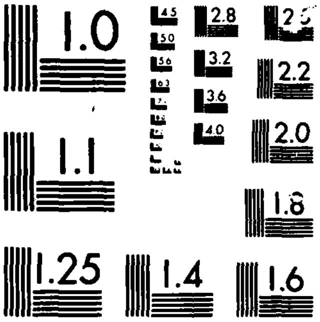
AD-A166 995   SEMICONDUCTOR ALLOY THEORY(U) AUBURN UNIV ALA DEPT OF   1/2
PHYSICS A CHEN 27 SEP 85 AFOSR-TR-86-0159
AFOSR-84-0282

UNCLASSIFIED   F/G 20/12   NL


1.0
4.5
5.0
5.6
2.8
3.2
3.6
4.0
2.2
2.0
1.1
1.8
1.25
1.4
1.6

MICROCOP CHART

AD-A166 995

(2)

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered) UNCLASSIFIED

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER **AFOSR-TR- 86-0159**
2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Semiconductor Alloy Theory
5. TYPE OF REPORT & PERIOD COVERED Annual 84-9-1 to 85-8-31
6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) An-Ban Chen
8. CONTRACT OR GRANT NUMBER(s) AFOSR-84-0282
9. PERFORMING ORGANIZATION NAME AND ADDRESS

Auburn University
AL, 36849

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2306/[redacted] B1
11. CONTROLLING OFFICE NAME AND ADDRESS

AFOSR/[redacted] NE
Building 410
Bolling AFB DC 20332

12. REPORT DATE 85-9-27
13. NUMBER OF PAGES 4
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)
15. SECURITY CLASS. (of this report) Unclassified
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report)

The United States Government is authorized to reproduce and distribute this report for government purpose

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

DTIC ELECTE APR 2 9 1986 S D D

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Electron Mobility, Lattice Relaxation, Bond Energies, Mixing Enthalpies, Hardness, Dislocation, Band Structure, Core-Exciton, Alloy Statistics, Microclustering, Deep Levels

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

See Reverse Side

DTIC FILE COPY

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

86 4 28 111

ABSTRACT

We have investigated many topics in the theory of semiconductors and alloys including (1) Generalized Brooks' formula and the electron mobility in SiGe Alloys, (2) bond-length distribution, lattice relaxation, bond energies and mixing enthalpies in alloys, (3) hardness and dislocation energy, (4) SiGe alloys - band structures and core-exciton binding energy and linewidth, (5) alloy statistics and microclustering and (6) sensitivity of deep levels to band structure and potential. The important results are outlined and the details are discussed in eight publications enclosed with the report. — 1473A

ANNUAL TECHNICAL REPORT

to

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
BOLLING AIR FORCE BASE, D. C. 20332

on

SEMICONDUCTOR ALLOY THEORY

(AFOSR-84-0282)

For Period

September 1, 1984 to August 31, 1985

Accession For
NTIS GRA&I ☑
DTIC TAB ☐
Unannounced ☐
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or Special
A-1

Submitted by

An-Ban Chen
Physics Department
Auburn University
AL 36849

QUALITY INSPECTED 3

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFSC)
NOTICE OF TRANSMITTAL TO DTIC
This technical report has been reviewed and is
approved for public release IAW AFR 190-12.
Distribution is unlimited.
MATTHEW J. KERPER
Chief, Technical Information Division

Summary

We have investigated many topics in the theory of semiconductors and alloys including (1) Generalized Brooks' formula and the electron mobility in SiGe alloys, (2) bond-length distribution, lattice relaxation, bond energies and mixing enthalpies in alloys, (3) hardness and dislocation energy, (4) SiGe alloys - band structures and core-exciton binding energy and linewidth, (5) alloy statistics and microclustering and (6) sensitivity of deep levels to band structure and potential. The important results are outlined below. The details are discussed in eight publications enclosed with this report.

Generalized Brooks' Formula and the Electron Mobility in SiGe Alloy

Although Brooks' formula has been widely used for calculating the alloy-scattering limited electron mobility, we showed that this formula is only valid for a direct-gap semiconductor. Besides, there are questions about the scattering parameter and the effective mass. We generalized the formula for indirect-gap alloys with multiple bands and applied it to SiGe alloy. Our results, correlated well with experiments, showed that the electron mobility drops fast with alloying. The mobility has a dip at 15% Si concentration, corresponding to a transition from the X(Δ) band edge to the L edge.

Bond Lengths, Lattice Relaxation, and Mixing Enthalpies in Semiconductor Alloys

We treated the problem with a model which combines Harrison's bonding theory with the valence force field model and an elastic continuum. While the local strain is the main driving force for the bimodal bond-length distribution in pseudo-binary alloys found in the EXAFS experiment, we found that the chemical shifts arising from different bond lengths and polarities of the constituent bonds can have a significant contribution to the mixing enthalpies. In fact, the sizable negative values of the chemical shifts in the cation-substitutional alloys, e.g. $Ga_xIn_{1-x}As$, may be important for stabilizing the mixture. We also deduced a simple criterion for separating miscible from immiscible alloys.

Hardness and Dislocation Energy

We showed that, if the dislocation energy in semiconductor is dominated by interactions among dislocations, then the hardness is an intrinsic material property, independent of the size of the indentation and the force applied. The fact that dislocation energy is proportional to hardness suggests that alloy hardening is a mechanism for reducing dislocation densities.

SiGe Alloys - Band Structure and Core-Excitons

SiGe binary alloy has regenerated research interests because of its potential application for high-mobility devices in the strained superlattice configuration. We have applied our techniques to obtain high-quality band structures for Si and Ge and have performed alloy calculations for Si_xGe_{1-x}. The results have been checked against optical spectra and used for mobility study mentioned earlier. Another interesting result is that the alloy band parameters allowed us to correlate the Si 2p core-exciton binding energy with the linewidth. The observed minimum in the linewidth near $X \simeq 0.15$ can be explained as the result of a competetion between intrinsic broadening due to screening and extrinsic alloy broadening. The most reasonable binding energy in pure Si was found to be 0.15 eV. This work thus helps resolve the controversy about the unusually large binding energy for the core exciton in Si.

Statistics and Micro-clustering in Semiconductor Alloys

One of our major efforts was to develop a statistical theory for semiconductor alloys which covers three aspects -energetics, statistics and phase diagram. Because of the smallness of the mixing enthalpies, typically several k Cal/mol, the present first-principle theory is not accurate enough for this purpose. We found that a combination of Harrison's model and the valence force field model provides a simple and adequate method for calculating the mixing energies. We then generalized Guggenheim's quasi-chemical approximation to treat the tetrahedral clusters. Our results showed that the arrangement of atoms can have a appreciable deviation from the random distribution. The distribution is governed by the mixing energy which is the sum of the strain energy and chemical shifts mentioned earlier. The non-random distribution will have profound effects on the band structures, mobilities, mechanical properties and growth mechanisms, and will reflect in many measurable properties such as phonon spectra, EXAFS, deep level spectra etc. The study of these effects is one of our current activities.

Publications (Enclosed with this report)

1. "Semiconductor Alloys: Local Bond Lengths, Mixing Enthalpies, and Microclusters", A.-B. Chen and A. Sher, paper presented at 1985 Materials Research Meeting, April 15-18, San Francisco, CA.

2. "Sensitivity of Defect Energy Levels to Host Band Structures and Impurity Potentials in CdTe", A.-B. Chen and A. Sher, Phys. Rev. B31, 6490 (1985).

3. "Effects Influencing the Structural Integrity of Semiconductors and Their Alloys", A. Sher, A.-B. Chen, W.E. Spicer, and C.-K. Shih, J. Voc. Sci. Technol. A3 (1), 105 (1985).

4. "Phenomena Influencing the Dislocation Density of Semiconductor Compounds and Alloys", A. Sher, A.-B. Chen and W.E. Spicer. Thirteen International Conference on Defects in Semiconductors, edited by L.C. Kimerling and J.M. Parsey, Jr. P. 335 (The Metallurgical Society of AIME, 1985).

5. "Dislocation Energies and Hardness of Semiconductors", A. Sher, A.-B. Chen and W.E. Spicer, Appl. Phys. Lett. 46, 54 (1985).

6. "Generalized Brooks' Formula and the Electron Mobility in $Si_xG\ell_{1-x}$ Alloys", S. Krishnamurphy, A. Sher and A.-B. Chen, Appl. Phys. Lett. 47, 160 (1985).

7. "Binding Energy and Spectral Width of Si 2p Core Exciton in $Si_xG\ell_{1-x}$ Alloys", S. Krishnamurphy, A. Sher and A.-B. Chen, Appl. Phys. Lett. 55, 320 (1985).

8. "Semiconductor Pseudo-Binary Alloys: Bond-Length Relaxation and Mixing Enthalpies", A.-B. Chen and A. Sher, Accepted for pulication in Phys. Rev. B.

Generalized Brooks' formula and the electron mobility in Si_xGe_{1-x} alloys

Srinivasan Krishnamurthy and A. Sher
SRI International, Menlo Park, California 94025

An-Ban Chen
Auburn University, Auburn, Alabama 36830

(Received 19 March 1985; accepted for publication 7 May 1985)

A formula for alloy-scattering-limited electron mobility in semiconductors is obtained for indirect gap systems with multiple band minima. All the input parameters needed are defined explicitly. The drift mobility of Si_xGe_{1-x} which has a dip at $x \sim 0.13$ and a broader minimum at $x \sim 0.5$ is calculated by adding alloy scattering to other scattering mechanisms and correlates well with the measured Hall mobility.

The electron and hole mobilities in semiconductors are determined by the band structure and various scattering mechanisms, predominately impurity and phonon scattering. For alloys, the mobility is also affected by disorder arising from aperiodic atomic potentials and atomic positions. Many years ago, Nordheim[1] and Brooks[2] obtained an expression for alloy-scattering-limited electron mobilities in metals and semiconductors, respectively. Brooks' well-known formula reads

$$\mu_A = \frac{\sqrt{2\pi} e \hbar^4 N_0}{3x(1-x) m^{*5/2}} \frac{1}{(\Delta E)^2 \sqrt{KT}}, \tag{1}$$

where N_0 is the number of atoms per unit volume, m^* is a band-edge effective mass, x is the fractional concentration of one of the species, and ΔE is an energy parameter characterizing the alloy potential fluctuations. Although this formula has been widely and, to some extent, successfully used for direct gap materials,[3-5] the identification of the alloy disorder parameter ΔE remains uncertain. Various suggestions have previously been made for ΔE, e.g., and band-edge discontinuity[6] or band-gap differences.[7] Any of these simple choices is bound to fail when one applies Eq. (1) to more complicated indirect gap systems such as Si_xGe_{1-x} alloys, where one encounters conduction-band minima transferring between the X and L points of the Brillouin zone. For example, if ΔE is taken to be the difference in corresponding band edges, then one finds that $\Delta E \sim 0.1$ eV for the X (Δ) valley and ~ 1.2 eV for the L valley. The values that fit the experiment are about half this value for L and ~ 0.5 eV for X.[6] The purpose of this letter is to resolve the identity of ΔE for indirect gap materials.

Moreover, there is a problem with the m^* that enters Eq. (1). For direct gap alloys, the band-edge effective mass at Γ naturally enters Eq. (1). For the indirect gap alloys, the effective mass is anisotropic and hence an appropriate mass must be chosen. Previous authors[6,7] have chosen m^* to be the effective conductivity mass m_c^*. We shall show that different masses enter for different cases.

The first unambiguous assignment for ΔE in a direct gap alloy was given by Hass *et al.*[5] To estimate the limiting electron mobility in $Hg_xCd_{1-x}Te$ based on a tight-binding (TB) band description, they defined ΔE to be $f_s \Delta E_s$, where f_s is the s fraction in the density of states and ΔE_s is the difference between the s atomic term values of the Hg and Cd atoms. By extending this approach to alloys with indirect gaps and multiple bands, we show that all the uncertainties identified above are resolved. Our generalized Brooks' formula will then be applied to Si_xGe_{1-x} systems to explain their observed mobility.[7,8]

Because Brooks' formula has never been derived explicitly in the literature, we rederive it first and then generalize it. Consider the case of a single band with an isotropic effective mass. The dc electronic conductivity based on the linear response theory[9] is given by

$$\sigma = \int \sigma(\epsilon) \left(\frac{-df}{d\epsilon} \right) d\epsilon, \tag{2}$$

where the energy-dependent $\sigma(\epsilon)$ in the weak alloy scattering limit is

$$\sigma(\epsilon) = (e^2/3) v^2(\epsilon) D(\epsilon) \tau(\epsilon). \tag{3}$$

$D(\epsilon)$ is the density of states (DOS) per unit volume for both spins, so $D(\epsilon) = 2N_A \rho(\epsilon)$, with $N_A = N_0/2$ being the number of unit cells per volume (for the diamond structure, half the number of atoms N_0 per unit volume; and $\rho(\epsilon)$ being the DOS per unit cell per spin. The mean square velocity $v^2(\epsilon)$ for carriers with energy ϵ is given by

$$v^2(\epsilon) = \sum_{\mathbf{k}} |v(\mathbf{k})|^2 \frac{\delta[\epsilon - \epsilon(\mathbf{k})]}{\rho(\epsilon)}. \tag{4}$$

The scattering lifetime for carriers with energy ϵ, $\tau(\epsilon)$, is related to the alloy broadening $\Delta(\epsilon)$ by $\tau(\epsilon) = \hbar/2\Delta(\epsilon)$, where the energy $\Delta(\epsilon)$ is the imaginary part of the self-energy in the averaged alloy Green's function. For weak scattering $\Delta(\epsilon)$ is

$$\Delta(\epsilon) = \pi x(1-x)(\Delta E)^2 \rho(\epsilon), \tag{5}$$

where in a tight binding (TB) description ΔE is the difference in the term values of the constituents. Then the mobility is $\mu_A = \sigma/ne$ with the electron density given by

$$n = 2\int f(\epsilon) \rho(\epsilon) d\epsilon. \tag{6}$$

For a nondegenerate semiconductor, $f(\epsilon)$ is the Boltzmann distribution and $f(\epsilon) \propto e^{-(\epsilon - \epsilon_f)/kT}$. Furthermore, for a parabolic band $\epsilon(\mathbf{k}) = \hbar^2 k^2/2m^*$, $\rho(\epsilon) = (2m^*)^{3/2} \epsilon^{1/2}/4\pi^2\hbar^3$, then all the above equations can be combined to arrive at Eq. (1).

For a real semiconductor alloy in a TB description, the alloy scattering can be characterized by two parameters ΔE_s and ΔE_p, the differences in s and p term values between two substitutional atoms. Then an effective broadening is given by

$$\Delta(\epsilon) = (\Delta_s\rho_s + \Delta_p\rho_p)/\rho, \qquad (7)$$

where ρ_s, ρ_p are partial density of states (PDOS) and Δ_s and Δ_p are similar to Eq. (5), with ρ replaced by ρ_s and ρ_p, respectively. For $Hg_{1-x}Cd_xTe$, the s disorder is predominent[5,10] and one can neglect ΔE_p. Defining $\rho_s = f_s\rho$ (and $\rho_p = f_p\rho$), one arrives at

$$\Delta \simeq \pi x(1-x)(f_s\Delta E_s)^2\rho.$$

Thus as was pointed out by Hass *et al.*,[5] $f_s\Delta E_s$ plays the role of ΔE in this special case where ΔE_p can be neglected.

For an alloy with a single indirect gap minimum, one has to consider both s and p contributions to the alloy broadening and the masses that enter ρ and v^2. Again, Eqs. (2)–(6) can be combined to yield

$$\mu_A = \frac{(e\hbar^4 N_0\sqrt{2\pi})}{[3x(1-x)m_c^* m_t^*(m_l^*)^{1/2}(kT)^{1/2}N_v(f_s^2\Delta E_s^2 + f_p^2\Delta E_p^2)]}, \qquad (8)$$

where m_l^* and m_t^* are respectively the longitudinal and the transverse mass at the band edge, and N_v is the number of equivalent minima, e.g., 6 for Si. The conductivity mass m_c^* comes from averaging v^2 in Eq. (2) and is given by $3(2/m_t^* + 1/m_l^*)^{-1}$. Equation (8) clearly identifies the masses and the energy parameter that enter Brooks' formula.

Next we consider a still more complicated case where the contribution to the mobility comes from more than one band. For example, in Si_xGe_{1-x} the X and L minima cross near $x = 0.15$.[11] There are now two contributions to the net conductivity, so $\sigma = \Sigma\sigma_i$, where i is X or L. The quantities $v_i^2(\epsilon)$, $D_i(\epsilon)$, and $N_i(\epsilon)$ now take different values for different bands. The structure of $\tau_i(\epsilon)$ requires more careful consideration. The complication comes from the fact that the effective broadening Δ is still given by Eq. (7), but ρ_s, ρ_p, and ρ contain contributions from both the bands. The proper expressions are $\rho = \Sigma_i\rho_i N_v^i$ and $\rho_\alpha = \Sigma_i f_{\alpha i}N_v^i\rho$, where $i = X$ or L, $\alpha = s$ or p, and $N_v^X = 6$, $N_v^L = 4$. The equation for Δ is

$$\Delta(\epsilon) = \pi x(1-x)\sum_\alpha\left(\sum_i f_{\alpha i}N_v^i\rho_i(\epsilon)\Delta E_\alpha\right)^2\bigg/\left(\sum_i \rho_i(\epsilon)N_v^i\right). \quad (9)$$

The mobility associated with the ith band is defined as $\mu_i = \sigma_i/(n_i e)$, then

$$\mu_i^A = \frac{\pi e\hbar^4 N_0}{3x(1-x)}\frac{1}{[m_c^* m_t^*(2m_l^*)^{1/2}]_i}I_i, \qquad (10)$$

$$I_i = \int_0^\infty \frac{\epsilon\rho_i(\epsilon)\left(\sum_j N_v^j\rho_j(\epsilon)\right)e^{-\epsilon/kT}}{\sum_\alpha\left(\sum_j f_{\alpha j}N_v^j\rho_j(\epsilon)\Delta E_\alpha\right)^2}d\epsilon\bigg/\int\epsilon^{1/2}e^{-\epsilon/kT}\,d\epsilon. \quad (11)$$

Thus, the generalized formula no longer has the explicit x and T dependences of the original Brooks' form. However, all the quantities needed—the masses, the scattering parameters ΔE_α, the band gaps, and the fractions $f_{\alpha i}$—can be evaluated theoretically without resorting to experimentally fitted parameters. To demonstrate, we shall apply Eq. (10) to Si_xGe_{1-x}. The band quantities are obtained from our CPA calculation.[11] We found that the effective masses vary weakly with the concentration, so m_l^* and m_t^* are assumed to be constant and assigned the values 0.97 and 0.19 for the X minima and 1.64 and 0.082 for the L minima, respectively. The calculated energy gaps for the $X(\Delta)$ follows the functional form $E_g^{(X)} = a + bx + cx^2$ and for L is given by $E_g^{(L)} = A + Bx + Cx^2$. All the parameters of our calculations are listed in Table I.

TABLE I. Calculation parameters.

Parameter	Si_xGe_{1-x} systems
$m_l^*(X)$	$0.97m_0$
$m_t^*(X)$	$0.19m_0$
$m_l^*(L)$	$1.64m_0$
$m_t^*(L)$	$0.082m_0$
$E_g^X(x)$	$0.8941 - 0.0421x + 0.1691x^2$
$E_g^L(x)$	$0.7596 + 1.0860x - 0.3306x^2$
$f_{sX}(x)$	$0.333 + 0.05x \quad (0\leqslant x\leqslant 0.3)$
	$0.339 + 0.03x \quad (0.3\leqslant x\leqslant 1.0)$
$f_{sL}(x)$	$0.632 + 0.13x$

To correlate the calculation with the measured mobilities, we need to have an estimate of scattering rates $1/\tau_0$ due to impurities and phonons. A crude approximation is to assume $1/\tau_0$ for a given valley to be the same as the appropriate constituent's values and add to it the alloy scattering rate $1/\tau_A$. Then the average mobility and the mobility from the ith minimum in the alloy are

$$\mu = \sum_i n_i\mu_i\Big/\sum_i n_i,$$
$$(\mu_i)^{-1} = (\mu_i^0)^{-1} + (\mu_i^A)^{-1}, \qquad (12)$$

μ_i^A is given by Eq. (10) and μ_i^0 are the measured drift mobilities for Si or Ge.[12] The drift mobility, calculated from Eq. (12), is plotted as a function of alloy concentration x in Fig. 1.

For $x<0.05$ and $x>0.20$, the energy difference between the X and L edges is large enough so there is a negligible contribution to the mobility from the higher minima. In the Si_xGe_{1-x} system, the s scattering is predominent. Because

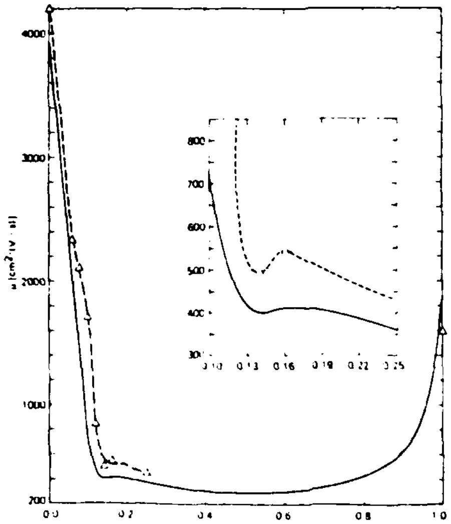


FIG. 1. Calculated (solid line) electron drift mobility and the experimental Hall mobility (dashed line) from Refs. 6 and 8 are plotted as a function of alloy concentration.

the L edges have a larger s content, alloy scattering reduces the average mobility substantially for small x. Even though the s content is almost the same for all $x \geq 0.20$ at the X edge, the mobility still decreases to $x \simeq 0.5$ as shown in Fig. 1 because of the $x(1-x)$ term in Eq. (10).

An interesting feature is obtained for the compositions $0.13 \leq x \leq 0.18$. The average mobility attains a local minimum near $x = 0.14$ and a smaller maximum at $x = 0.17$. This feature occurs because of the X to L crossover.[11] For $x < 0.14$, the major contribution to μ comes from L minima. Near $x = 0.14$, the density of states increases because the X and L minima merge. So the alloy scattering increases there and the average mobility decreases. For $x \geq 0.14$, the X bands have the lower minima. As the s content is small at the X minima, the reduced alloy scattering increases the average mobility. For larger values of x, the $x(1-x)$ term takes over and the mobility varies as shown. The values of measured Hall mobility in $\mathrm{Si}_x\mathrm{Ge}_{1-x}$ systems are also plotted in Fig. 1. The interesting feature near $x = 0.14$ is clearly seen. Since the experimental drift mobility μ_D for $\mathrm{Si}_x\mathrm{Ge}_{1-x}$ is not available and the generalization of Eq. (10) to Hall mobility μ_H is less clear, we present the calculated μ_D and experimental μ_H (Ref. 7,8) here. While we do not expect quantitative agreement, because μ_H/μ_D can range from 1 to 2,[13,14] we do expect them to display the same qualitative x dependence. It is rewarding to note the similarity in the trend in Fig. 1. Previous authors explained the dip in the mobility curve by including intervalley scattering with an arbitrary adjustable coupling constant.[6] Our calculations automatically include that portion of intervalley scattering that results from alloy disorder with a coupling constant set by the atomic properties of the constituents. However, the additon of intervalley scattering mediated by phonons and impurities is expected to increase the dip near crossover.

In summary, an expression for alloy-scattering-limited charge carrier mobilities is derived for indirect gap alloys with multiple bands. This expression reduces to Brooks' formula for direct gap alloys. The quantities m^* and ΔE can be calculated exactly. Alloy scattering accounts for the observed mobility features in the $\mathrm{Si}_x\mathrm{Ge}_{1-x}$ alloy, including the anomaly near the L to $X(\Delta)$ crossover.

A.-B. would like to thank Professor W. E. Spicer for his hospitality at Stanford University. This work was supported in part by DARPA contract MDA 903-83-C-0108 and grant AFOSR-84-0282.

[1] L. Nordheim, Ann. Phys. **9**, 607 (1931); **9**, 641 (1931).

[2] H. Brooks (unpublished). A discussion of this formula can be found, for example, in L. Makowski and M. Glicksman, J. Phys. Chem. Solids **34**, 487 (1973).

[3] A. Chandra and L. F. Eastman, J. Appl. Phys. **51**, 2669 (1980).

[4] D. Chattopadhyay and B. R. Nag, Phys. Rev. B **12**, 5676 (1975).

[5] K. C. Hass, H. Ehrenreich, and B. Velicky, Phys. Rev. B **27**, 1088 (1983).

[6] J. W. Harrison and J. R. Hauser, J. Appl. Phys. **47**, 292 (1976).

[7] M. Glicksman, Phys. Rev. **111**, 125 (1958).

[8] M. Glicksman, Phys. Rev. **100**, 1146 (1955).

[9] A.-B. Chen, G. Weisz, and A. Sher, Phys. Rev. B **5**, 2897 (1972). See Eq. (134).

[10] D. S. Montgomery, J. Phys. C **16**, 2923 (1983).

[11] S. Krishnamurthy, A.-B. Chen, and A. Sher (unpublished).

[12] S. M. Sze, *Physics of Semiconductors*, 2nd ed. (Wiley–Interscience, New York, 1981).

[13] V. A. Johnson and K. L. Horowitz, Phys. Rev. **79**, 176 (1950); **79**, 409 (1950).

[14] H. Jones, Phys. Rev. **81**, 149 (1951).

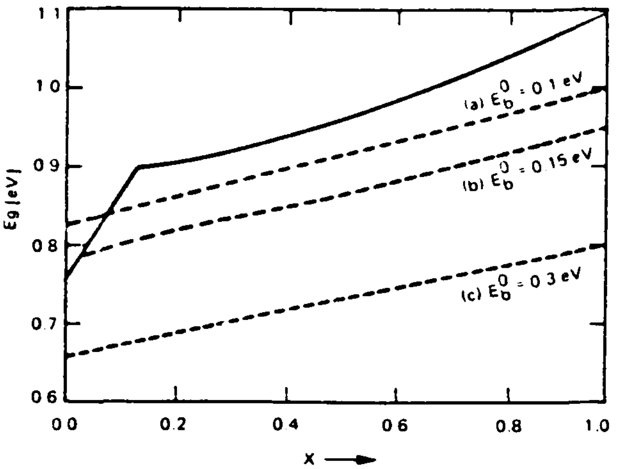


FIG. 1. Variation of the band gap (solid line) and the Si $2p$ core-exciton level with x in $\mathrm{Si}_x\mathrm{Ge}_{1-x}$ alloys. The energy is measured from the top of the valence band. The dashed curves a, b, and c represent exciton levels calculated with $E_b^0 = 0.1$, 0.15, and 0.3 eV, respectively.

by replacing the short-range Coulomb potential with a spherical square well of variable depth and a screened Coulomb tail. Strinati's results can be used to estimate Δ_l corresponding to the calculated E_b. Δ_l decreases rapidly with E_b, then saturates for larger E_b.

The contribution to the natural linewidth from the alloy broadening is calculated by a consideration of the electron part of the exciton wave function, ψ_s. The ψ_s is expanded in a linear combination of the s part of the conduction-band wave functions ϕ_n^s:

$$\psi_s(\mathbf{k}) = \sum_{n,s} C_{ns}\phi_n^s(\mathbf{k}). \tag{7}$$

We found that alloy scattering is only moderate and s scattering is dominant; thus, the alloy broadening $\Delta_A(E)$ is well approximated by

$$\Delta_A(E) \approx x(1-x)\delta_s^2 \,\mathrm{Im} F_s(E), \tag{8}$$

where δ_s is the difference between E_s^{Si} and E_s^{Ge}. Hence, the alloy-broadening contribution to Δ is related to the alloy broadening of the band states, $\Delta_n(\mathbf{k},\mathbf{E})$:

$$\begin{aligned}\Delta_A &= \frac{1}{N}\sum_{\mathbf{k}} \langle \psi_s(\mathbf{k}) | \Delta_A(E) | \psi_s(\mathbf{k})\rangle \\ &= \frac{1}{N}\sum_{\mathbf{k}}\sum_{n} C_{n,s}^2 \Delta_n(\mathbf{k},E) \\ &= \int \rho_s(E)\Delta_A(E)\,dE \\ &= x(1-x)\delta_s^2\pi \int \rho_s^2(E)\,dE. \end{aligned} \tag{9}$$

The integral in Eq. (9) is evaluated numerically.

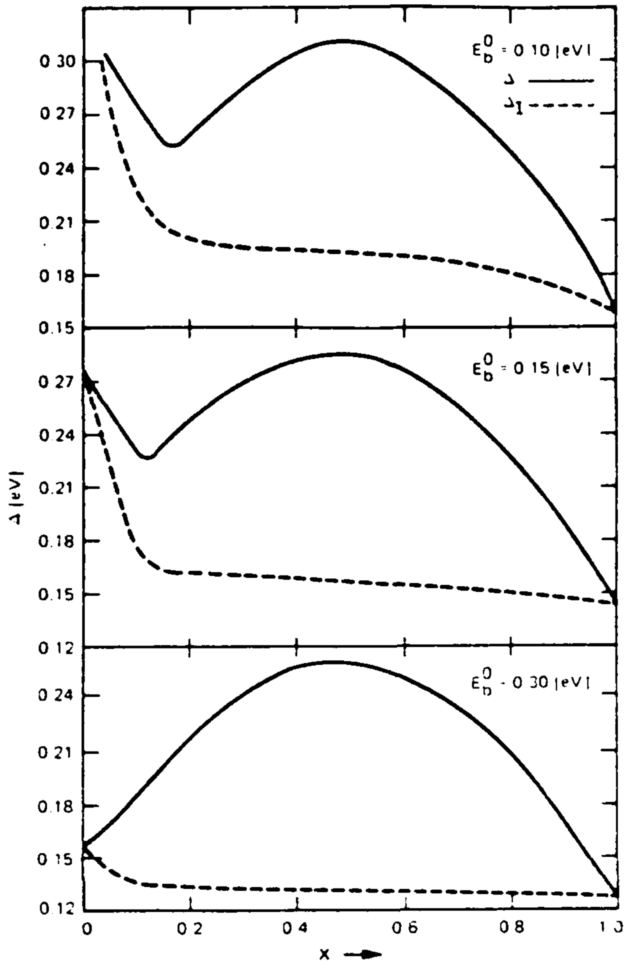


FIG. 2. Variation of Δ (solid lines) and Δ_l (dashed lines) with x for three E_b^0 values.

The calculated Δ, which is the sum of Δ_l and Δ_A, is plotted against x in Fig. 2 for three values of E_b^0. In all three panels, the dashed curve represents Δ_l and the solid line represents Δ. It is seen from Fig. 1 that the exciton level follows the X edge of the conduction band. Hence the binding energy E_b, relative to the conduction band edge, remains almost constant (for a given E_b^0) until the minimum switches from the X edge to the L edge. Because of the change in the slope of E_g, E_b decreases rapidly when L becomes the minimum. Correspondingly, Δ_l varies slowly until the X to L crossover and then increases rapidly. This feature is clearly seen in Fig. 2.

For $E_b^0 = 0.15$, the Δ_l and Δ_A are comparable near $x = 0.50$, and Δ_l dominates for all small x and large x. These two competing mechanisms give a relative minimum near $x \approx 0.15$, a broader maximum near $x \approx 0.50$, and a smaller minimum for pure silicon. As E_b^0 is decreased, the relative minimum is shifted to larger x, e.g., the minimum shifts to $x = 0.20$ for $E_b^0 = 0.10$ eV. For $E_b^0 = 0.15$ eV, the position of the

Edelman, Solid State Commun. **36**, 298 (1980).

[7] C. Kunz, J. Phys. (Paris), Colloq. **39**, C4-119 (1978).

[8] H. P. Hjalmarson, H. Büttner, and J. D. Dow, Phys. Rev. B **24**, 6010 (1981).

[9] M. Altarelli, Phys. Rev. Lett. **46**, 205 (1981).

[10] G. Strinati, Phys. Rev. Lett. **49**, 1519 (1982).

[11] K. E. Newman and J. D. Dow, Solid State Commun. **50**, 587 (1984).

[12] B. A. Bunker, S. L. Hulbert, J. P. Stott, and F. C. Brown, Phys. Rev. Lett. **53**, 2157 (1984).

[13] S. E. Schnatterly and R. D. Carson, Bull. Am. Phys. Soc. **30**, 416 (1985).

Dislocation energies and hardness of semiconductors

A. Sher
SRI International, Menlo Park, California 94025
A.-B. Chen
Auburn University, Auburn, Alabama 36849
W. E. Spicer
Stanford University, Stanford, California 95305

(Received 13 July 1984; accepted for publication 5 October 1984)

The dislocation energies and hardness of semiconductors are calculated by an extension of Harrison's method. It is demonstrated in agreement with experiment that dislocation energies per unit length are proportional to $d^{-3}-d^{-9}$, where d is the bond length and hardness is proportional to $d^{-5}-d^{-11}$. The hardness is related to the interaction energies among dislocations. It is argued that dislocation densities of semiconductors will be reduced if they are alloyed with a second constituent that has a shorter bond length. Experimental evidence supporting this strategy is noted.

Dislocations in semiconductors are detrimental to device function; they serve as channels for impurity migration and trapping, which cause nonuniform doping and degrades p-n junctions.[1] They also decrease the material's resistance to plastic deformation. The aim of this letter is to provide insights into the underlying physical mechanisms controlling dislocations and semiconductor hardness, and then to suggest strategies for decreasing dislocation densities. It is well established that the hardness of tetrahedrally coordinated semiconductor materials—groups IV, III-V, and II-VI compounds—exhibits a sharp variation with their near-neighbor distance d, approximately proportional to d^{-9} for one group of seven compounds.[2] Thus, semiconductors with small lattice constants tend to be harder materials. These same materials have larger stiffness coefficients[3] and have fewer dislocations in as-grown crystals.[4–6]

The shear coefficients (combinations of C_{44} and $C_{11}-C_{12}$ in the Schoenflies notation)[7] depend on crystal orientation and (in Harrison's notation[8]) are proportional to $V_2^3/d^5(V_2^2+V_3^2)^{3/2}$, where $V_2 \propto d^{-2}$ is the covalent and V_3 is the ionic energy. The metallic interaction modifies the functional dependence of the shear coefficient on V_2 and V_3, but introduces no explicit dependence on the hopping integrals, denoted V_1 by Harrison.[3,8] In a pure covalent material, the bond energy is proportional to V_2 (or d^{-2}), and the bond volume is $\propto d^3$; hence, in this case, the shear coefficient varies as d^{-5}. In the limit, $V_3 > V_2$, $C_{11}-C_{12} \propto d^{-11}$. For most polar semiconductors, d^{-9} is a good approximation.

Hardness is determined by applying a known force F to a probe of a prescribed shape driving it into the surface of the sample.[7] The area A of the resulting indentation is measured, and the hardness is the force per unit indented area. Many dislocations must be formed to allow the probe to indent the semiconductor. If the indenter is a rectangular pyramid, then the hardness is $H = F/A = Fh/Ah = \epsilon_T/Ah$, where ϵ_T is the work required to cause the indenter to penetrate to a depth h. A side view of the indentation in a cut through its center is illustrated schematically in Fig. 1. The top of the indentation has side length W; thus, $A = W^2$. The Burger's vector has magnitude b, proportional to the bond length d. The number of dislocations N_h required to accommodate an indentation to depth h is $N_h = h/b = \frac{1}{2}W\cos\vartheta/b$, where ϑ is the angle between the normal to the tip of the indenter and a side. Figure 1 also shows a model of one possible configuration of the dislocations. The edges of the extra atom planes that are driven from the indented volume into the bulk of the semiconductor are shown as lines terminated by dots. The dotted ends of these lines are the positions of the dislocations, which are perpendicular to the plane of the figure. The planes driven to the sides each have a finite extent and a trapezoidal shape. The planes driven down under the indenter have a square shape.

Much of the work done on the indenter goes into the energy to form the indicated dislocation configuration, although some certainly goes into heat. There are two major contributions to this formation energy. The first is the enegy needed to generate each dislocation as an isolated entity, and the second is the interaction energy among these dislocations. Because the interaction term dominates H, approximations made to simplify the first term are relatively unimportant. The extra planes driven to the sides of the indentation have a finite extent; accordingly (in this idealized picture), there are both edge dislocations at their base and screw dislocations associated with their termination. The square planes driven below the indentation have edge dislocations around the sides and screw dislocations at the corners to make the turns. Moreover, there are interactions

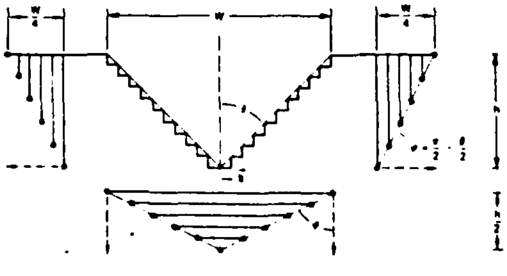


FIG. 1. Schematic representation of an idealized minimum-energy dislocation configuration produced by a square cross-section indenter in a hardness measurement. W is the side length and h is the depth of the indentation. In this ideal case, half the material from the indented region is displaced along the glide planes (indicated by the dashed arrows) to the sides and half is displaced below the indenter.

© 1985 American Institute of Physics

among the dislocations, which can produce a minimum-energy configuration. For the arrangement depicted in Fig. 1, it always costs energy to position a second dislocation on a parallel glide plane to one already present. However, the magnitude of this extra energy can be minimized and, for proper configurations of the dislocations, there are attractive forces along the glide planes that will tend to position the dislocations into the minimum-energy configuration. The minimum configuration arises when half of the atoms from the indented volume go respectively to the side and below the indenter. Then, in both regions, the maximum angle made between successive close-spaced dislocation lines and their glide planes is $\vartheta = \pi/2 - \vartheta/2$, as shown in the figure. This is the minimum realistic energy configuration. If the dislocations are separated more than shown in Fig. 1, then there is more volume of strained material and the interaction energy would be larger still.

An approximate expression for the energy required to indent the material is[7]

$$\epsilon_T = 2\cdot4\left[\sum_{i=1}^{N} E_i L_i + \frac{1}{2}\sum_{i,j=1}^{N}{}' E_{ij}\left[\min(L_i,L_j)\right]\right] \qquad (1)$$

In the first term, using an isotropic medium approximation and neglecting core terms, the energy per unit length to form an edge dislocation is[7]

$$E_i = \frac{Gb^2}{4\pi(1-\nu)}\ln\frac{R}{r_0}, \qquad (2)$$

the shear coefficient is G, the Burger's vector $b = d/3$ for an indentation along a $\langle 100\rangle$ axis, the range of the elastic deformation of a dislocation R is taken equal to W (for want of a better approximation), $r_0 \sim d$ is the dislocation core radius, the Poisson ratio is $\nu \sim 0.2$ for most semiconductors, and L_i is the length of the ith dislocation. In the second term, E_{ij} is the interaction energy per unit length between dislocations i and j. Assuming they have parallel glide planes and their Burger's vectors have the same sign, E_{ij} is given by[7]

$$E_{ij} = \frac{Gb^2}{2\pi(1-\nu)}\left[\ln\left(\frac{R}{r_{ij}}\right) + \cos^2\varphi_{ij}\right], \qquad (3)$$

where r_{ij} is the separation, and φ_{ij} is the angle that a line perpendicular to and joining the dislocations makes with the glide plane, as shown in Fig. 1. Because the various dislocations in a region have different lengths, the net interaction energy is approximated by multiplying the energy per unit length by the length of the shorter one. The upper limits on the sums N are the number of dislocations in one region (side or bottom) associated with one edge. For the minimum-energy configuration, $N = N_h/2$. The four that multiplies the bracket accounts for the four sides, and the two for the two regions for each side.

We now encounter our first surprise. As we can see from Eq. (2) and (3), E_{ij} and E_i have comparable magnitudes. Because there are approximately N^2 terms in the interaction energy sum, only N terms in the formation energy sum, and N is a large number ($N \gg 1$), the interaction energy completely dominates the hardness. In fact, N is typically of the order 10^5. Hence, terms owing to screw dislocations, core energies, heat dissipation as the dislocations propagate to their places, and other effects associated with the first term are unimportant. However, care must be taken with the interaction terms: Eq. (1) neglects a number of secondary interactions, some positive and others negative; these will be added later. The principal neglected terms are the interactions between the dislocations in the different regions on each side (positive), and the interactions between adjacent side and bottom regions (positive), and the interaction between the opposite sides in the bottom regions (negative). Comparison of the results with experiment will indicate how important these neglected terms are likely to be. The length L_i is $W(N-i)/N$ for i from 1 to N. This is the largest length of the side-inserted planes and its choice partially accounts for interacitons between the otherwise neglected screw dislocations. The distance $r_{ij} = r_i - r_j$ is given by $2\sqrt{2b}\,|i-j|$ for i and j ranging from 1 to N for the minimum-energy configuration and a tetrahedrally bonded semiconductor. Finally, in the indicated configuration, $\varphi_{ij} = \varphi = \pi/2 - \vartheta/2$. Inserting these expressions into Eq. (1) and retaining only terms of order N^2 yields

$$H_{\min} = \frac{G\cot\vartheta}{6\pi(1-\nu)}\left[-\ln\left(\frac{\cot\vartheta}{\sqrt{2}}\right) + \frac{4}{3} + \sin^2\frac{\vartheta}{2}\right]. \qquad (4)$$

One can also get a number for the hardness of a dislocation in which all the material is pushed along the same glide plane, e.g., to the side, to the bottom, or normal to the face of the indenter (a possibility not depicted in Fig. 1). In this case, the factor of 2 in front of Eq. (1) is removed, $N = N_h$ and $\phi = \pi/2 - \vartheta$. Then a higher nonequilibrium hardness in the context of this model (denoted H_1), is obtained.

$$H_1 = \frac{G\cot\vartheta}{3\pi(1-\nu)}\left[-\ln\left(\frac{\cot\vartheta}{\sqrt{2}}\right) + \frac{4}{3} + \sin^2\vartheta\right]. \qquad (5)$$

The proper answer for most materials, and depending on crystal orientation, probably lies somewhere between $H_{\min}$ and H_1. For an indenter with $\vartheta = \pi/4$, we have $H_{\min} = 0.0969G/(1-\nu)$ and $H_1/H_{\min} = 2.39$. Harrison[8] has shown that one contribution to the shear coefficient (actually $C_{11} - C_{12}$) is $G = 2.38\ \hbar^2 m\alpha_c^3/md^5$, where m is the free-electron mass, α_c is the covalence, $\alpha_c = V_2/(V_2^2 + V_3^2)^{1/2}$, and d is the bond length. We will approximate G by this expression. Using this G and $\nu = 0.2$, and changing the dimensions to those in terms of which experimental hardness numbers are customarily quoted gives $H_{\min} = 2.38\times 10^4\,(\alpha_c^3/d^5)$ kgm/mm^2, where d is in angstroms. Calculated values of $H_{\min}$ and H_1 are plotted against experimental results in Fig. 2 for a number of semiconductors.

Figure 2 has the theoretical $H_{\min}$ and H_1 values connected by arrows from $H_{\min}$ to H_1 for each compound, plotted as a function of the corresponding experimental values.[2,9] If the theory were perfect and the experimental values were accurate, the points would fall on the indicated unity slope line. Several conclusions can be drawn. Firstly, the order of magnitude of the predicted and measured values are the same, a result obtained with no adjustable parameters in the theory. Secondly, the trends from one compound to another are properly given by the theory. Although the $H_{\min}$ values are generally too small, they fit the soft materials better, and the H_1 values fit the harder materials better. Thirdly, from Eqs. (1) and (3), H is given in a rough but revealing

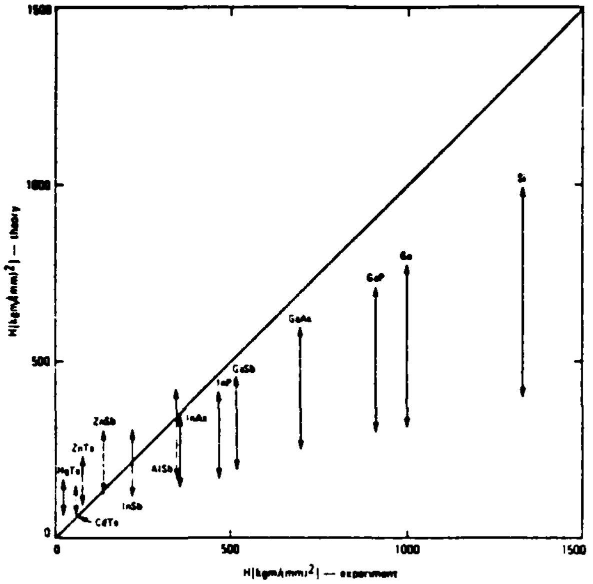


FIG. 2. Theoretical vs experimental hardness of several semiconductors in Refs. 2 and 9. The two theoretical values for each semiconductor are connected by an arrow from H_{min} to H_1. Perfect agreement would correspond to points being on the unity-slope solid line.

approximation by $4N^2\bar{E}_{ij}\bar{L}/Ah$, where the averge dislocation length $\bar{L} = W/2$ and $\bar{E}_{ij}$ is the average dislocation pair interaction energy per unit length. Notice that N (or W) cancels from this expression; thus, H is independent of W (or F) and therefore, H is truly a measure of the properties of the material. This result would not be found if the dislocation energies ($\propto N$) were to dominate H rather than the pair interaction energy ($\propto N^2$). Finally, the Berger's vector cancels from the leading term and appears only in the argument of the logarithm in Eq. (3). Thus, the answers are also insensitive to its choice.

Dislocations are often found in materials as they are grown. Their density is determined by the thermal and mechanical stresses to which they are subjected in the growth process. A dislocation constitutes a metastable excitation relative to the perfect crystal ground state. At the elevated growth temperatures and temperature gradient behind the growth front, the number of dislocations present is controlled by the relative rate at which vacancies anneal or condense into dislocations.[10] The dislocation formation rate will be slower in a material grown at the same temperature if E_l is higher. If an alloy is formed from a material of interest and a second constituent with a shorter bond length, one expects the average bond energy (and thus both the melt temperature and vacancy formation energy) to increase proportional to a low inverse power of the average bond length.[8] Hence, the equilibrium vacancy density just below the melting point tends to be the same in lowest order for all materials, independent of the bond lengths of the constituents. However, the shear coefficient and dislocation energy per unit length will increase with much higher inverse powers of the bond length. Consequently, dislocation densities should be reduced in such alloys relative to those found in the longer bond length pure constituent. This expectation is confirmed in the recent work reported on $Zn_{1-x}Cd_xTe$ bulk material.[4] The best CdTe that has ever been grown has dislocation densities in excess of 5×10^5 cm^{-2}. The addition of only 4% Zn reduced the dislocation count to less than 5×10^4 cm^{-2}. The ZnTe bond length is 2.643 Å, while the CdTe bond length is 2.805 Å, a 6% difference. This 6% difference in bond length translates into a 2% difference in the dislocation energy per unit length for $1 - x = 0.04$. Dislocation energies per unit length are typically 10 eV per lattice spacing; accordingly, a 2% increase can be expected to slow their formation rate considerably.

The argument just presented naturally leads to a strategy for decreasing dislocations in other semiconductors. If an alloy is made of the material of interest with another compound with a shorter bond length, then the dislocation density should be reduced. For example, this suggests that the addition of a small amount of GaP ($d = 2.359$ Å) may significantly reduce the dislocation density of bulk grown GaAs ($d = 2.448$ Å). It has been demonstrated that the addition of approximately 1% GaN ($d = 1.946$ Å)[5,6] or of a 10^{18} cm^{-3} BAs ($d = 2.069$ Å) concentration[11] to GaAs can yield a large volume of dislocation-free material. An InAs additive with its longer bond length ($d = 2.623$ Å) serves the same function indirectly, by causing GaAs bonds in its neighborhood to be compressed. This indirect mechanism should be less effective than substituting short bond length additives.

We have demonstrated that the dislocation energies and hardness of tetrahedrally bonded semiconductors are rapid functions of the reciprocal of the bond length. This rapid d dependence of dislocation energies provides a rationale for the dramatic decrease of the dislocation density in bulk grown $Zn_{0.04}Cd_{0.96}Te$ material relative to that found in CdTe, and suggest means for accomplishing the same ends in other materials.

The authors are indebted to J. P. Hirth, W. A. Harrison, and T. N. Casselman for helpful comments. This work was supported in part by DARPA contract MDA 903-83-C-0108 and AFOSR contract 49620-81-K-0012.

---

[1] E. M. Swiggard, Proceedings of GaAs IC Symposium, Phoenix, AZ 1983, p. 26.

[2] N. A. Groyunova, A. S. Borshchevskii, and D. N. Fretiakov, in *Semiconductors and Semimetals*, edited by R. K. Willardson and A. C. Beer (Academic, NY, 1968), Vol. 4, Chap. 1.

[3] W. A. Harrison, *Electronic Structure and the Properties of Solids* (Freeman, San Francisco, 1980); R. C. Sokel, thesis, Stanford University, 1978; W. A. Harrison, *Microscience* (limited distribution SRI International publication, Menlo Park, 1983), Vol. 4, p. 34.

[4] S. L. Bell and S. Sen, presented at Infrared Imaging Systems (IRIS) Detector Specialty Group Meeting, Boulder CO, 1983; T. W. James and B. F. Zuck, *ibid.*

[5] Y. Seki, H. Watanabe, and J. Matsui, J. Appl. Phys. **42**, 822 (1983).

[6] G. Jacob, J. Cryst. Growth 59, 669 (1982).

[7] J. W. Christian, *Theory of Transformations in Metals and Alloys*, 2nd ed. (Pergamon, NY, 1975), Chap. 7. For a discussion of the interaction energies, see J. P. Hirth and J. Lothe, Theory of Dislocations, 2nd ed. (Wiley, NY, 1982), pp. 262–263.

[8] W. A. Harrison, Phys. Rev. B **27**, 3592 (1983).

[9] S. Cole and A. F. W. Willoughby, J. Cryst. Growth **59**, 370 (1982).

[10] G. Schoeck and W. A. Tiller, Philos. Mag. **5**, 43 (1960).

[11] S. Miyazawa, 1983 European Patent Application 833021652, filing date 4/18/83, and references therein.

PHYSICAL REVIEW B VOLUME 31, NUMBER 10 15 MAY 1985

Sensitivity of defect energy levels to host band structures and impurity potentials in CdTe

A.-B. Chen
Physics Department, Auburn University, Auburn, Alabama 36849

A. Sher
SRI International, 333 Ravenswood Avenue, Menlo Park, California 94025

(Received 10 December 1984)

The sensitivity of defect energy levels in semiconductors to the host band structures and impurity potentials has been studied for approximately 30 impurities in CdTe using four different band-structure models. The discrepancies in the defect levels between two different sets of band structures and impurity potentials are found to range from less than 0.1 eV to the whole band gap (1.6 eV). The band-structure effects are analyzed here in terms of detailed partial densities of states. Examples of contradictory predictions from different band structures are illustrated, and ways to improve the theory are suggested.

I. INTRODUCTION

In several of our recent papers,[1-7] we have applied a method to calculate the band structure of semiconductors that is both efficient and accurate. Because the procedure involves casting the basis functions into orthonormal local orbitals[6] (OLO), our method has the advantages common to empirical tight-binding (ETB) calculations,[8-10] except that the Hamiltonian matrix elements to all ranges are retained. The inclusion of these higher coefficients makes it possible to produce excellent band structures including conduction bands and effective masses. The method also yields wave functions for optical property calculations.[7] Moreover, its OLO description also permits its extension, through the coherent-potential approximation, to alloys.[2-5]

The recent attention focused on defects in semiconductors motivated us to apply our method to this problem. The theories of defects have ranged from very sophisticated self-consistent density-functional theory[11-13] (SCDF) to simple ETB calculations. It is generally recognized that SCDF is as accurate in defects for the ground-state properties as it is for pure semiconductors, but less certain in assigning excited energy levels. ETB, because it can produce results for many systems in one study, claims to predict the trends of deep levels[10] even if the accuracy for a given impurity may be poor. However, this contention remains to be verified.

To assess this concern, we ask the following question: "How sensitive are defect levels to host band structures and impurity potentials?" To this end, we have adopted the simple yet nontrivial defect model, that of site-diagonal substitutional defects often used in ETB studies. CdTe was selected in this study because its band structure has been examined in great detail by us, and there are three published band-structure models[8-10] that we could easily generate for comparison. There is also a considerable body of experimental data on deep states in this system.[14-17]

II. CALCULATIONAL PROCEDURE

In the simple site-diagonal substitutional defect model, the impurity energy levels E are determined by the equation

$$1-v_\alpha g_\alpha(E)=0\ , \tag{1}$$

where α designates the symmetry of a local state, e.g., Γ_6, Γ_7, and Γ_8 on an atomic site in the zinc-blende structure, and g_α is the real part of the diagonal matrix element of the host-crystal Green function. g_α can be calculated from the partial density of states (PDOS) by

$$g_\alpha(E)=\int \rho_\alpha(\epsilon)/(E-\epsilon)d\epsilon\ . \tag{2}$$

The PDOS is given by

$$\rho_\alpha(\epsilon)=\sum_{n,k} |\, a_n^\alpha(k)\,|^2\delta(\epsilon-\epsilon_n(k))\ , \tag{3}$$

where $\epsilon_n(k)$ are band energies and $a_n^\alpha(k)$ are the probability amplitudes of the band state in the Bloch basis constructed from the OLO labeled by α. The Brillouin-zone integration in Eq. (3) is calculated using an accurate ray scheme.[18]

Because a principal concern of this paper is the sensitivity of impurity levels to the host band structures, we should emphasize the difference between our method and ETB. Our method consists of four steps.

(1) We start with four Gaussian orbitals per atom and empirical pseudopotentials,[19] and compute the Hamiltonian matrix $H(k)$ and overlap matrix $S(k)$ as was done by Kane[20] and Chadi.[21]

(2) The Gaussian orbitals are transformed into OLO,[6] so $H(k)$ is transformed into $H_0(k)$ and S into the identity matrix. The band structures calculated from $H_0(k)$ are accurate to 5% as compared to more sophisticated methods using the same potential.[1]

(3) A spin-orbit Hamiltonian in the OLO basis[4] is incorporated to deal with this interaction.

(4) To compensate for the effects of truncated basis and

©1985 The American Physical Society

nonlocal potentials, a perturbation Hamiltonian H_1 is added. H_1 has the same form as a truncated ETB Hamiltonian. The parameters in H_1 are adjusted to fine tune the important band energies and effective masses.[1–4]

Although both ETB and our methods are empirical, there are two major differences.

(1) While most ETB retains the H matrix elements only to the first- or second-neighbor shell, ours extends to all ranges, so that the high Fourier components needed to produce the sharp band curvatures are properly given.

(2) Our method can directly generate wave functions for calculation of other properties.

Thus, while our method yields more accurate band structures, it retains much of the advantage of ETB, namely the computational speed and a simple direct-space description of the Hamiltonian.

III. BAND STRUCTURES AND PARTIAL DENSITIES OF STATES

Figure 1 depicts the four band structures to be considered for CdTe. Our result is in panel (a); panels (b) (Ref. 8) and (c) (Ref. 9) are two ETB band structures with the Hamiltonian matrix elements truncated at second neighbors. (Because different parameters were selected, these two band structures are not identical.) Panel (d) (Ref. 10) results from the use of five basis orbitals per atom; the extra one is an excited s state. All these band structures are adjusted to have the proper fundamental band gap of 1.6 eV. The principal differences one sees on first inspection are in the band curvatures, especially the conduction bands. The effective mass at the bottom of the conduction band in panel (a) is 0.1 times the free-electron mass, in agreement with experiment,[17] while in other panels it is more than twice as large.

Figure 2 shows the densities of states (DOS) for each of the band structures in Fig. 1. While the valence bands at least exhibit general common features, the conduction bands are almost unrecognizable as representing the same compound. In panels (c) and (d), for example, there is a second band gap above the fundamental gap. Also note that there are two extra narrow peaks associated with the two extra excited s orbitals (one for Cd and the other for Te) included in the calculation.

To analyze the band effects on defect levels [see Eqs. (1) and (2)], the DOS is further decomposed into partial densities of states for $\Gamma_6(s)$, $\Gamma_7(p^{1/2})$, and $\Gamma_8(p^{3/2})$ states on the Cd and Te sites, as shown in Figs. 3–6. The Γ_8 PDOS are not shown because they are nearly the same as Γ_7 with only a slight upward energy shift. These PDOS show how the "atomic" levels evolve into band states. These curves contain useful information about many properties, e.g., the relation between the crystal bonding and atomic energies, and how potential disorder in alloys affects different parts of the bands,[2–5] in addition to defect levels studied here.

The Γ_6(Cd) PDOS shown in Fig. 3 split between the conduction and valence bands. It is generally assumed that the cation s states in III-V and II-VI compounds

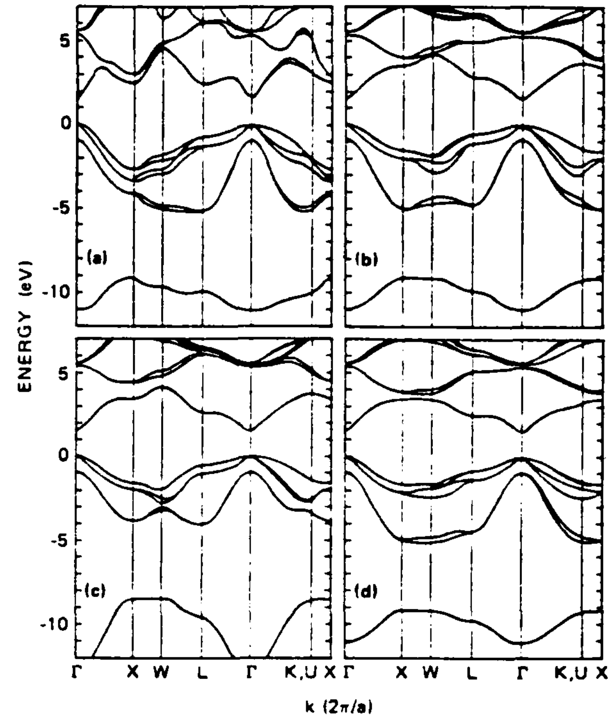


FIG. 1. Four band structures of CdTe used for comparative studies: (a) present work, (b) Ref. 8, (c) Ref. 9, and (d) Ref. 10.

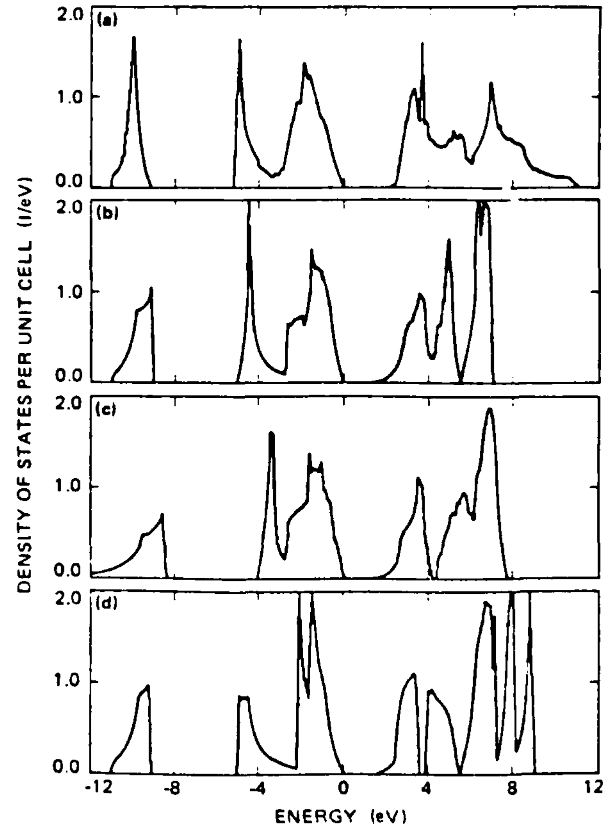


FIG. 2. Densities of states calculated from the four band structures in Fig. 1.

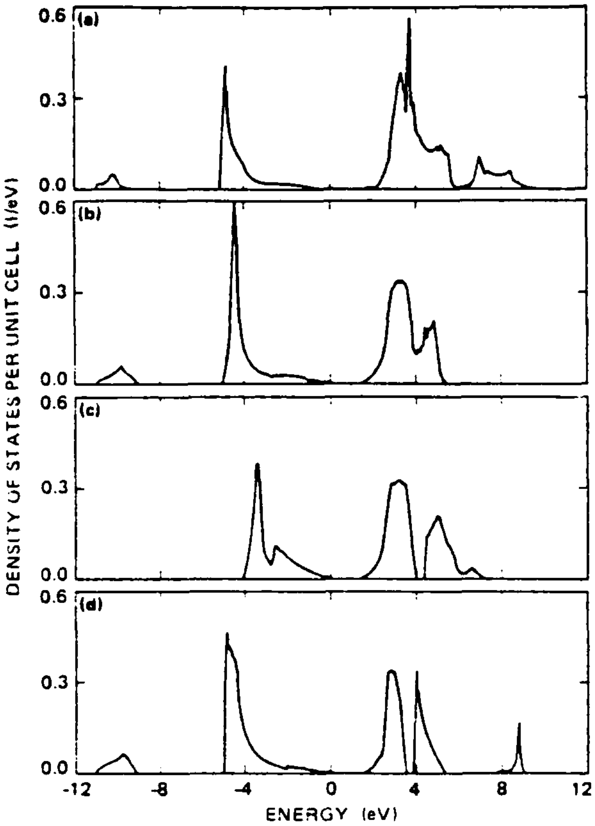


FIG. 3. The Cd Γ_6 partial densities of states.

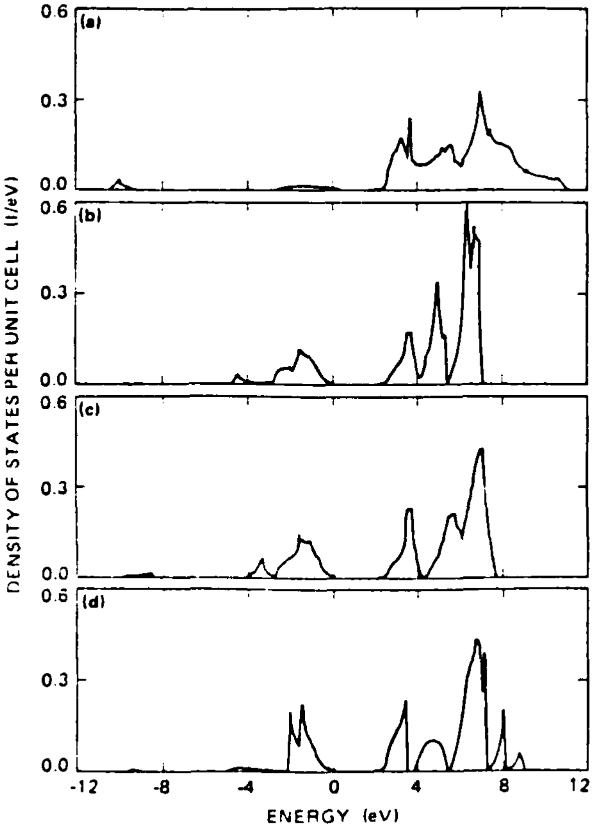


FIG. 4. The Cd Γ_7 partial densities of states.

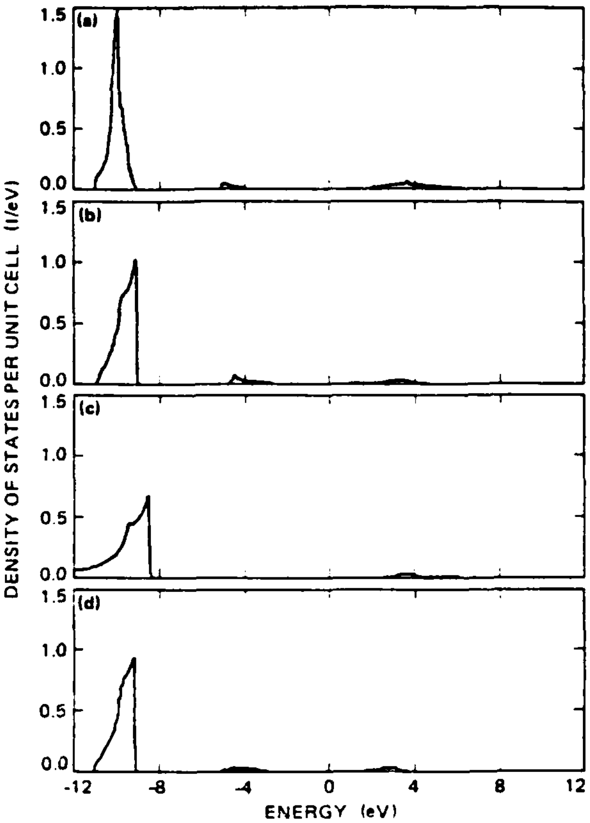


FIG. 5. The Te Γ_6 partial densities of states.

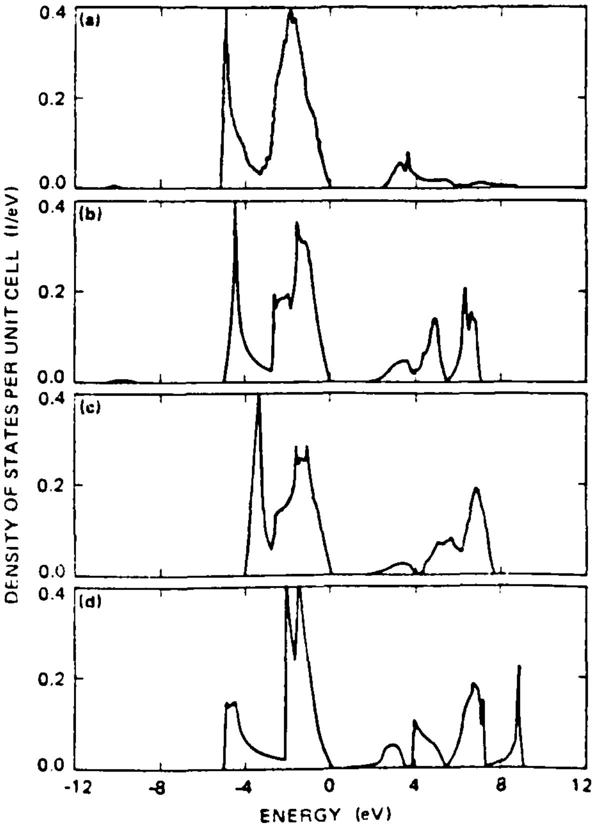


FIG. 6. The Te Γ_7 partial densities of states.

evolve into the conduction bands, while the anion *p* states make up most of the major valence bands just below the gap. Thus it is perhaps a surprise to see a prominent peak derived from the cation *s* states at the bottom of the major valence-band structure. However, this is a general feature for all sp^3-based compound semiconductors. These are the states responsible for the first observed breakdown of the virtual-crystal approximation for a semiconductor alloy: $Hg_{1-x}Cd_xTe$ (which is caused by the large *s*-energy shift between the Cd and Hg sites).[4,5,22]

A more detailed examination draws attention to some important differences among the four panels in Fig. 3: the valence-band peak in panel (c) is about 2 eV higher than the rest, and it is also high compared to experiment.[22] Our conduction-band PDOS in panel (a) is broader than the others. The ratio of the integrated PDOS in the conduction bands to that in the valence bands in our model is larger than those in other panels. Also our PDOS just below the valence-band edge is obviously smaller than that found in other models.

Figure 4 shows that the Cd *p* states are concentrated in the conduction-band states. This is particularly true in panel (a), where their contribution to the valence-band states shrinks almost to nothing. In other panels, there are still sizable (~20%) valence-band states. In contrast, all four panels in Fig. 5 show that the Te *s* states are confined to the deep valence-band states, as generally recognized. Finally, Fig. 6 shows that the Te *p* states dominate the upper valence-band states. Panel (a) has much less conduction-band content than the other three panels. As we will see, these differences can result in quantitatively or even qualitatively different predictions about the deep levels.

IV. IMPURITY-LEVEL DETERMINATION

A convenient way to study the impurity energy levels using Eq. (1) is to rewrite it as $v_\alpha = 1/g_\alpha(E)$ and plot E as a function of v. Once this E-v curve is deduced for each α, the deep levels E_α for a given impurity can be read off the curve by drawing a vertical line at the appropriate value of v_α for the impurity. We set the zero of energy at the top of the valence bands. Because the gap is 1.6 eV, we will focus on levels in the energy range from −0.5 to 2.0 eV.

Calculations have been performed for all neutral impurities listed in Table I. Because we do not believe that there exists a uniformly accepted table for v we have adopted a table that we used for structural studies.[23,24] Table I lists the term values, which we obtained from to-

TABLE I. *s*- and *p*-state correlated term values in units of −eV. The top entry is the *s*-state, the second the $p_{1/2}$-state, and the third the $p_{3/2}$-state energy. (All energies are *negative*.)

I	II	III	IV	V	VI	VII
Li	Be	B	C	N	O	F
5.390	9.320	14.003	19.814	26.081	28.551	36.229
	5.412	8.300	11.260	14.540	13.613	17.484
	5.412	8.300	11.260	14.540	13.610	17.420
Na	Zn	Al	Si	P	S	Cl
5.140	9.390	11.780	15.027	19.620	21.163	25.812
	4.237	5.980	8.150	10.610	10.449	13.136
	4.011	5.980	8.150	10.550	10.360	13.010
K	Cd	Ga	Ge	As	Se	Br
4.340	8.990	13.230	16.396	20.015	21.412	24.949
	4.313	6.000	7.880	10.146	10.188	12.353
	4.097	5.850	7.694	9.810	9.750	11.840
Rb	Hg	In	Sn	Sb	Te	I
4.180	10.430	12.032	14.525	17.560	19.120	21.631
	4.998	5.780	7.340	9.391	9.951	11.470
	4.031	5.453	6.879	8.640	9.010	10.450
Cs			Pb			
3.890			15.250			
			7.410			
			5.979			
Cu						
7.720						
Ag						
7.570						
3.647						
3.487						
Au						
9.220						
4.349						
3.688						

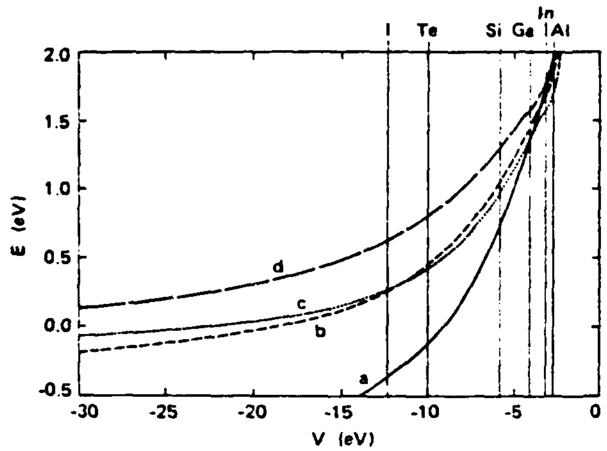


FIG. 7. The E-v curves for the Γ_6 states on a Cd site.

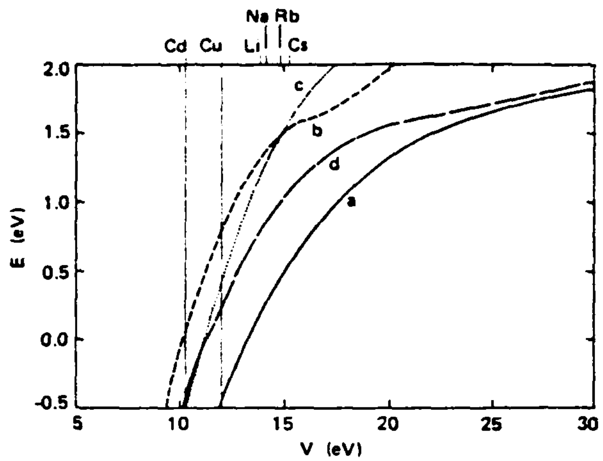


FIG. 9. The E-v curves for the Γ_6 states on a Te site.

tal energy differences between atomic configurations calculated using the norm-conserved pseudopotentials[25] and self-consistent charge-density-functional theory, with the first ionization energies adjusted to be the experimental values.[26] These term values are found to yield consistently better structural properties[23] in Harrison's theory[27,28] than those based on Mann's values[29] adopted by Harrison.[28] The impurity-potential parameters will then be taken as the difference of the term values between the impurity atom and Cd (or Te). To study the sensitivity of E_α to v_α, we shift v_α by ±0.5 eV and compute the corresponding changes in the energy levels.

Figures 7–10 display the E-v curves for several α. Each figure has four curves, corresponding to the four panels of PDOS in each of Figs. 3–6. The functional behavior of these curves can be understood qualitatively using Eq. (2) and Figs. 3–6. If E lies in the gap, the contribution from conduction bands is negative, but positive from the valence bands. The closer the PDOS to the E in question, the larger will be its influence. Applying this argument to the Γ_6(Cd) representation, we see that the curves in Fig. 7 are negative in the gap region because the PDOS in Fig. 3 near the bottom of the conduction bands are much larger than those near the valence-band top. Thus, on the Cd site, only impurities with an s energy below the Cd s level (−8.99 eV) will produce a Γ_6 level in the gap. However, we note that in Fig. 7, $g_\alpha(E)=0$ for models (b) and (c) just below the valence-band edge because of cancellation between the conduction- and valence-band contributions. At this E value, the E-v curve switches from $v=-\infty$ to $v=\infty$ (not shown); an ideal vacancy level (corresponding to $v_\alpha=\infty$) is located at this E. A similar consideration, but with the conduction and valence bands interchanged, leads to an understanding of the curves in Fig. 10. Using the same principle, we can easily understand why all curves in Fig. 9 for the Γ_6(Te) representation are positive, but the reasons for the large displacements between these curves are not easy to deduce. In Fig. 8, the curve labeled a is distinctively different from other curves, because the PDOS in panel (a) in Fig. 4 is completely dominated by the conduction band; however, for the other panels the PDOS just below the valence-band edge are as large as those just above the conduction-band edge. This produces a very sharp negative E-v curve for (a), but split behavior for (b), (c), and (d).

These E-v curves provide a clear picture of how different host band structures may affect the deep levels. Numerical values for the impurity levels can be obtained from these figures by drawing vertical lines at the ap-

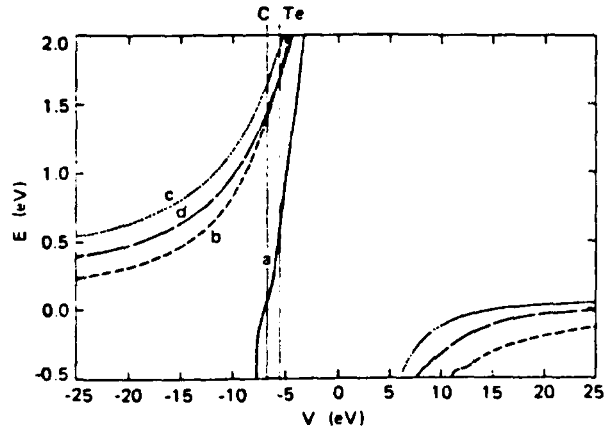


FIG. 8. The E-v curves for the Γ_7 states on a Cd site.

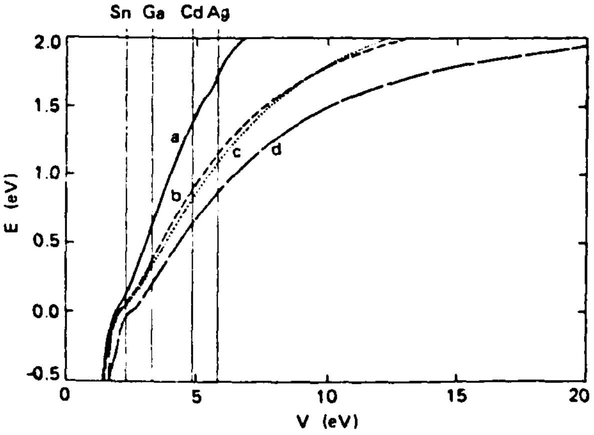


FIG. 10. The E-v curves for the Γ_7 states on a Te site.

propriate impurity potentials (i.e., differences between the term values listed in Table I), as has been shown for several representative impurities. To provide a more quantitative comparison, Table II lists some calculated impurity levels E_α and the corresponding changes ΔE_α due to the 1-eV change in v_α.

V. RESULTS AND CONCLUSION

To summarize we recall that band models (b) and (c) are the same second-neighbor ETB with two different sets of parameters, and model (d) is a first-neighbor ETB with one extra s orbital per atom. Our model [model (a)] has the form of ETB but is derived in a very different manner and includes all the long-range interactions. Therefore, we expect that the results from models (b) and (c) will be close, model (d) will have larger discrepancies from (b) and (c) than that between (b) and (c), and model (a) will differ even more. This is evident from Figs. 7–10 and Table II. We found the energies for the Γ_6(Cd), Γ_7(Te), and Γ_8(Te) states produced by models (b) and (c) agree within 0.1 eV. For the other states, i.e., Γ_6(Te), Γ_7(Cd), and Γ_8(Cd), the energies from (b) and (c) are qualitatively similar, but the difference can be as large as 0.4 eV. The largest discrepancy between models (d) and (b) [or (c)] is more than 0.5 eV, and that between (a) and other models is more than 1 eV. The largest difference comes from the p levels on a Cd site. For example, the filled p level of C on a Cd site in model (a) is a resonance state just below the valence-band edge but is a donor state in the other models. Similarly, model (a) puts the neutral Te antisite defect p levels at about $\frac{1}{3}$ and $\frac{2}{3}$ of the gap [$E(\Gamma_7)=0.48$ eV and $E(\Gamma_8)=0.95$ eV], while other models assign them as resonance states inside the conduction bands. We also note that the discrepancies between different models are not uniform, but vary with v_α. Consider Γ_6(Cd) for example. All four models yield the same ordering and about the same energies for the group-III impurities Al, In, and Ga. However, as v becomes more negative, the splitting between the curves increases, so the discrepancies become larger [~ 1 eV difference between models (a) and (d) for I impurity]. Similarly, for the Γ_7(Te) states, all four models put the Sn impurity energies close to the valence-band edge, but the agreement deteriorates as v_α increases.

Regarding the sensitivity of energy levels to impurity potentials, Table II shows that a 1-eV shift in v_α produces a change in E_α ranging from less than 0.1 to 0.65 eV. Very little is known about the size or trends in errors introduced in v_α from the use of atomic term values. However, we know that the discrepancy of v_α between two different tables of atomic term values can be larger than 2 eV. This discrepancy translates into an uncertainty of less than 0.1 to more than 1 eV in the impurity energy levels,

TABLE II. Defect energy levels E and changes ΔE due to a 1-eV change in the impurity-potential parameter. All energies are in units of eV. V_0 stands for ideal vacancy.

Defect	Model (a) E	Model (a) ΔE	Model (b) E	Model (b) ΔE	Model (c) E	Model (c) ΔE	Model (d) E	Model (d) ΔE
Γ_6 on Cd site								
Ga	1.29	0.39	1.42	0.24	1.33	0.23	1.57	0.18
C	−0.21	0.09	0.38	0.09	0.36	0.13	0.74	0.08
Si	0.67	0.30	1.02	0.10	0.93	0.19	1.27	0.15
P	−0.19	0.11	0.39	0.09	0.38	0.08	0.75	0.08
O	<-0.5		−0.02	0.02	0.04	0.01	0.32	0.02
Te	−0.13	0.13	0.44	0.10	0.42	0.08	0.79	0.09
Cl	<-0.5		0.06	0.03	0.10	0.02	0.41	0.04
V_0	<-0.5		<-0.5		−0.30		−0.20	
Γ_7 on Cd site								
C	−0.02	0.37	1.32	0.22	1.59	0.20	1.39	0.19
Si	1.57	0.65	>2.0		>2.0		>2.0	
P	0.16	0.38	1.48	0.26	1.73	0.23	1.52	0.21
O	<-0.5		0.89	0.14	1.22	0.13	1.03	0.12
Te	0.48	0.55	1.60	0.29	1.88	0.23	1.66	0.24
Cl	<-0.5		0.96	0.17	1.29	0.14	1.09	0.14
V_0	<-0.5		0.00		0.21		0.06	
Γ_6 on Te site								
Li	0.14	0.29	1.28	0.22	1.15	0.35	0.76	0.25
Cu	<-0.5		0.54	0.42	0.12	0.52	0.03	0.32
Γ_7 on Te site								
Ag	1.89	0.32	1.26	0.22	1.21	0.23	0.99	0.20
Cd	1.66	0.34	1.11	0.26	1.05	0.26	0.85	0.22
Ga	0.98	0.49	0.61	0.33	0.55	0.32	0.40	0.30
Si	−0.07	0.40	−0.11	0.36	−0.13	0.38	−0.38	0.72
Sn	0.28	0.47	0.15	0.31	0.13	0.24	0.02	0.28

which is comparable to that due to different host band structures.

Putting this large uncertainty in the deep levels against a band gap of 1.6 eV, we are left with great doubts about the predictability of this oversimplified theory. Unfortunately, the experimental means available for identifying microdefects in semiconductors are still very limited, and the *ab initio* band theory is still not capable of accurately predicting the energy levels. Thus, there is a great temptation to use simple theories like the one carried out here to help with the identifications. To illustrate this point, consider the following examples: Table II shows that Li on a Te site has an s level of 0.14 eV in model (a), so one may be tempted to relate it to the acceptor state identified experimentally.[14] However, this is not the hydrogenic acceptor state on a Cd site, as one might anticipate. One might also want to assign the $\frac{1}{3}$ and $\frac{2}{3}$ gap states for the Te antisite p levels on the Cd site found from model (a) as those seen in experiments.[15,16] Because of the large uncertainty in the calculation, these results should be regarded as suspicious surprises rather than theoretical confirmations.

The results presented here should not discourage continued research on the ETB approach, but improvement is clearly needed. Work ranging from universal[23,27,30] to specific[24,31,32] structural studies to our band calculations and alloy studies[1-7] indicates that the ETB type of theory is practical for both bonding properties and electronic structures. The reason that ETB works well for some properties, e.g., photoemission spectra and bonding properties, but not for impurity levels, is that the former depend only on the gross total density of states, while the latter have been shown to be sensitive to the details of the partial densities of states.

To establish the credibility of ETB in defect studies, one needs to look at the problem more seriously. The most difficult and yet important task is to develop a better way for determining the Hamiltonian matrix elements. Haas *et al.*[8] and Harrison[27,28] have suggested using the atomic term values as the diagonal matrix elements. Our work[1-4] has suggested using a universal long-range interaction to improve the accuracy of the conduction bands. Several studies[1,27,28,33] have also pointed out scaling rules of the matrix elements. A combination of these ideas may lead to an acceptable model. Secondly, both the bonding and deep-level states of impurities should be studied at the same time in order to provide correlated information for defect identification. Finally, more realistic models should be examined. Besides the substitutional site-diagonal defects, one should consider the possibility of interstitial, paired, and even more complex defects. One also needs to deal with long-range impurity potentials, possible charge shifts, and lattice distortions. Progress in all these areas can be expected if the calculation is constantly correlated with experiments and available *ab initio* theory.

ACKNOWLEDGMENTS

This work was supported by AFOSR Contract No. F49620-81-K0012 and Grant No. AFOSR-84-0282. A.-B. Chen would like to thank Professor W. E. Spicer of Stanford University for his hospitality.

---

[1] A.-B. Chen and A. Sher, Phys. Rev. B **22**, 3886 (1980).

[2] A.-B. Chen and A. Sher, Phys. Rev. B **23**, 5645 (1981).

[3] A.-B. Chen and A. Sher, Phys. Rev. B **23**, 5360 (1981).

[4] A.-B. Chen and A. Sher, J. Vac. Sci. Technol. **121**, 138 (1982).

[5] W. E. Spicer, J. A. Silberman, J. Morgan, J. Lindau, J. A. Wilson, A.-B. Chen, and A. Sher, Phys. Rev. Lett. **149**, 948 (1982).

[6] A.-B. Chen and A. Sher, Phys. Rev. B **26**, 6603 (1982).

[7] A.-B. Chen, S. Phokaichaipatana, and A. Sher, Phys. Rev. B **23**, 5360 (1981).

[8] K. C. Hass, H. Ehrenreich, and B. Velicky, Phys. Rev. B **27**, 5360 (1983).

[9] C. A. Swarts, M. S. Daw, and T. C. McGill, J. Vac. Sci. Technol. **121**, 199 (1982).

[10] A. Kobayashi, O. F. Sankey, and J. D. Dow, Phys. Rev. B **25**, 6367 (1982).

[11] J. Bernholc, N. O. Lipari, and S. T. Pantelides, Phys. Rev. B **21**, 3545 (1980).

[12] G. A. Baraff, E. O. Kane, and M. Schluter, Phys. Rev. B **21**, 5662 (1980).

[13] G. B. Bachelet, M. Schluter, and G. A. Baraff, Phys. Rev. B **27**, 2545 (1983).

[14] K. Zanio, in *Semiconductors and Semimetals*, edited by R. Willardson and A. C. Beer (Academic, New York, 1978), Vol. 13.

[15] C. E. Jones, V. Nair, J. Lundquist, and D. L. Polla, J. Vac. Sci. Technol. **21**, 187 (1982).

[16] R. T. Collins and T. C. McGill, J. Vac. Sci. Technol. **A1**, 1633 (1983).

[17] K. K. Kanazawa and F. C. Brown, Phys. Rev. **135**, A1757 (1964).

[18] A.-B. Chen, Phys. Rev. B **16**, 3291 (1977).

[19] D. J. Chadi, J. P. Walter, and M. L. Cohen, Phys. Rev. B **5**, 3058 (1972).

[20] E. O. Kane, Phys. Rev. B **13**, 3478 (1976).

[21] D. J. Chadi, Phys. Rev. B **16**, 3572 (1977).

[22] J. A. Silberman, P. Morgan, W. E. Spicer, and J. A. Wilson, J. Vac. Sci. Technol. **21**, 142 (1982).

[23] A.-B. Chen and A. Sher, Microscience **3**, 1 (1984).

[24] A. Sher, A.-B Chen, and W. E. Spicer, in *Thirteenth International Conference on Defects in Semiconductors*, edited by L. C. Kimerling and J. M. Parsey, Jr. (The Metallurgical Society of AIME, 1985), p. 335.

[25] G. B. Bachelet, D. Hamman, and M. Schulter, Phys. Rev. B **26**, 4199 (1982).

[26] C. Kittel, *Introduction to Solid State Physics, 5th edition* (Wiley, New York, 1976), p. 75, Table 2.

[27] W. A. Harrison, *Electronic Structure and Properties of Solids*

(Freeman, San Francisco, 1980).

[28]W. A. Harrison, Microscience 3, 35 (1983).

[29]J. B. Mann, *Atomic Structure Calculations, I: Hartree-Fock Energy Results for Elements Hydrogen to Lawrencium* (Clearinghouse for Technical Information, Springfield, Virginia, 1967).

[30]D. G. Pettifor and R. Podloucky, Phys. Rev. Lett. 53, 826 (1984).

[31]D. C. Allan and E. J. Mele, Phys. Rev. Lett. 53, 826 (1984).

[32]D. J. Chadi, Phys. Rev. Lett. 52, 1911 (1984).

[33]O. K. Andersen, W. Klase, and M. Nohl, Phys. Rev. B 17, 1209 (1978).

Effects influencing the structural integrity of semiconductors and their alloys[a]

Arden Sher
SRI International, Menlo Park, California 94025
An-Ban Chen
Auburn University, Auburn, Alabama 36489
W. E. Spicer[b] and C-K Shih
Stanford University, Stanford, California 94305

(Received 1 June 1984; accepted 21 August 1984)

The bond length and energy changes of the constituents of alloys relative to their pure crystal values are calculated from an extension of Harrison's method. It is demonstrated that the already weak HgTe bonds are destabilized by adjacent CdTe, HgS, or HgSe, but are stabilized by ZnTe. It is also argued that dislocation energies and the hardness of semiconductors vary as a high inverse power of the bond length of the constituents. Hence, the shorter ZnTe bond as an additive should improve the structural properties of HgTe and CdTe. Experiments that support these predictions are noted. The electronic transport properties of 0.1 eV band gap HgZnTe are about the same as those of HgCdTe, and the structural properties of the Zn compound are superior; thus, we conclude that HgZnTe is likely to be the better material for IR devices.

I. INTRODUCTION

The objective of this work is to understand the microscopic mechanisms that govern the stability of $Hg_{1-x}Cd_xTe$ alloys, and then to suggest changes in the material that improve its strength without adversely affecting its electronic behavior. In pursuit of this goal we have extended Harrison's bonding theory,[1] which is applicable to all tetrahedral structured semiconductors, to calculate bond length and energy changes in an alloy—including charge shift and reconstruction effects—relative to their pure crystal values, vacancy formation energies, dislocation energies, and hardness.

In this paper, we review the experimental situation in several of these areas and compare some of the results with theory for all the group IV, III–V compound, and II–VI compound semiconductors and their alloys. For example, the theory properly predicts the observed inverse ninth power bond length (d^{-9}) dependence of the hardness[2] of semiconductors and, with no adjustable parameters, their correct magnitude. We demonstrate that the weak HgTe bond is destabilized by alloying it with CdTe, HgS, or HgSe; however, the bond is stabilized by ZnTe. Moreover, because the bond length of ZnTe (2.406 Å) is 14% shorter than that of HgTe (2.797 Å) or CdTe (2.804 Å), the dislocation energy per unit length and hardness of the alloys $Hg_{1-x}Zn_xTe$ and $Cd_{1-x}Zn_xTe$ are predicted in agreement with some experiments to be significantly higher than those of the compounds with $x = 0$. Measurements indicate that the electron and hole mobilities of $Hg_{1-x}Zn_xTe$ with $x \sim 0.16$, corresponding to a 0.1 eV band gap, are comparable to those of $Hg_{1-x}Cd_xTe$ for $x = 0.2$.[3] This is to be expected, because the electron effective mass in a narrow-gap material is small; hence, the electron wave functions at Γ_{1C} are distributed over many atoms, with the result that alloy scattering rates are small at the band edge. The hole mobilities are large for a somewhat different reason: The valence band edge is mostly composed of Te p states; however, there is also some cation p state contribution. Because the p state energies of Hg and Zn are nearly the same, the alloy scattering strength at the valence band edge is small. We conclude that, because of its structural advantage, HgZnTe may be superior to HgCdTe for infrared devices

II. ALLOY MODIFICATION OF THE BOND ENERGY AND LENGTH

A. Pure compounds

The first task is to recast Harrison's bond energy and bond length formalism into a structure suitable for generalization to alloys. Focus attention on the nth bond of a pure zinc blende structured compound semiconductor; for the time being, suppress any notation identifying it. Then, in a II–VI compound, the bond energy is[1]:

$$E_b = \frac{\epsilon_{p+} - \epsilon_{p-}}{2} + |V_{1+} + V_{1-}| - 2(V_2^2 + V_3^2)^{1/2} + 2\sum_{i=0}^{3}(\epsilon'_{m-} + \epsilon'_{m+}) + \frac{\gamma}{d^4}, \tag{1}$$

where the first term is the energy per bond needed to transfer two electrons from a p state on the group VI atom (the anion) to a p state on the group II atom (the cation), so that both start with four electrons (ultimately in the final bonding arrangement there is a net electron transfer from the cation to the anion), the second term is the promotion energy per bond to form sp^3 hybrids on both atoms, the third term is the bond formation energy owing to the covalent and ionic terms, the fourth term is Harrison's metallization energy, and the fifth term is the repulsive overlap energy that prevents the bonds from collapsing. The various symbols are defined below:

$\epsilon_{p(\pm)}$ is the $\left\{\begin{smallmatrix}\text{cation}\\ \text{anion}\end{smallmatrix}\right\}$ p state correlated term values,[4]
$V_{1(\pm)} \equiv (\epsilon_{p(\pm)} - \epsilon_{s(\pm)})/4$ are hopping integrals between two adjacent bonds coupled through a $\left\{\begin{smallmatrix}\text{cation}\\ \text{anion}\end{smallmatrix}\right\}$,
$\epsilon_{s(\pm)}$ are the $\left\{\begin{smallmatrix}\text{cation}\\ \text{anion}\end{smallmatrix}\right\}$ s state correlated term values,
$V_2 \equiv -24.5/d^2$ (eV) is the covalent energy,
d is the bond length,

0734-2101/85/010105-07$01.00
© 1985 American Vacuum Society

$$V_3 \equiv (\epsilon_{h+} - \epsilon_{h-})/2,$$

$\epsilon_{h\pm} \equiv (\epsilon_{s\pm} + 3\epsilon_{p\pm})/4$ are the sp^3 hybrid energies.

$$\epsilon^o_{m i\pm} \equiv \sum_{i=1}^{3} (U^o_{b\pm} U^i_{a\pm} V^2_{1i\pm})/(\epsilon^o_b - \epsilon^i_a)$$

is the contribution to the reduction of the bonding state energy ϵ^o_b calculated by second-order perturbation theory of the bond in question (denoted by superscript o), owing to its interaction with the three neighboring ($i = 1$–3) antibonding states sharing the same $\left\{\begin{smallmatrix}\text{cation}\\ \text{anion}\end{smallmatrix}\right\}$

$$U_{b+} = \sqrt{(1 - \alpha_p)/2} = U_{a-},$$

$$U_{b-} = \sqrt{(1 + \alpha_p/2} = U_{a+},$$

are the probability amplitudes of finding the cation and anion sp^3-hybrid contributions to the bonding and antibonding states, e.g., U_{b+} is the probability amplitude of finding the cation ^{3} hybrid in the bonding state, etc.,

$\alpha_p \equiv V_3/\sqrt{V_2^2 + V_3^2}$ is the polarity,

$\alpha_c \equiv V_2/\sqrt{V_2^2 + V_3^2}$ is the covalency,

$$\epsilon_{\substack{b\\a}} = \frac{\epsilon_+ + \epsilon_-}{2} (\mp)\sqrt{V_2^2 + V_3^2},$$

are the bonding and antibonding one electron energy states owing only to covalent and ionic interactions,

$$\epsilon_{mi\pm} \equiv \frac{(U^i_{b\pm} U^o_{a\pm} V_{1i\pm})^2}{\epsilon_b - \epsilon^o_a}, \quad i = 1,2,3$$

are the contributions to the bond in question arising from its antibonding state interactions with a neighboring bonding state i, and γ is an adjustable coefficient chosen so the measured bond lengths are reproduced.

Harrison has written Eq. (1) without explicit reference to the surrounding bonds. Equation (1) in its present form is rather easily used for alloy calculations by replacing host atom terms by one appropriate to the alloy.

The parameter γ for each compound is fixed by the condition that there is no net force on atoms at their equilibrium lattice positions. This condition is satisfied if the gradient of the total energy with respect to all the bond lengths vanishes:

$$\nabla_{\{d_n\}} E_T = 0, \quad \text{where } E_T = \sum_n E^n_b.$$

The condition is equivalent to the requirement that the net bond tension $T_{on} + T_{mn}$ for the nth bond equals the repulsive force

$$T_{on} + T_{mn} - \frac{4\gamma_n}{d^5_n} = 0, \tag{2}$$

where

$$T_{on} = \frac{\partial}{\partial d_n}(-2\sqrt{V_2^2 + V_3^2}) = 4\alpha_c\sqrt{V_2^2 + V_3^2}/d, \tag{3}$$

$$T_{mn} = 2\frac{\partial}{\partial d_n}\sum_{i=0}^{3}(\epsilon^i_{m+,n} + \epsilon^i_{m-,n}),$$

$$= \sum_{i=0}^{3}(T^i_{m+,n} + T^i_{m-,n}). \tag{4}$$

TABLE I. Correlated term values and bond energies for three II-VI compounds. The experimental bond lengths, polarity, and metallic contribution to the bond energy are also plotted.

Atom	Correlated term values (eV) s	p	sp^3
Zn	−10.224	−4.920	−6.246
Cd	−9.611	−4.784	−5.991
Hg	−10.946	−4.872	−6.391
Te	−19.620	−9.824	−12.273

Compound	Bond energies d (Å)^a	α_p	E_{metallic} (eV)	E_b (eV)
ZnTe	2.643	0.652	−0.625	−0.985
CdTe	2.805	0.710	−0.466	−0.944
HgTe	2.797	0.685	−0.614	−0.480

^{a} Reference 12.

The last identity in Eq. (4) defines the indicated bond tensions. Collecting all the results from the earlier paragraphs, one finds (dropping the bond designation n):

$$T^o_{m\pm} = -\frac{4}{d}\left[(1 \pm \alpha_p)\alpha_p \pm \frac{\alpha_c^2\sqrt{V_2^2 + V_3^2}}{\epsilon^o_b - \epsilon^1_a}\right]\epsilon^o_{m\pm}, \tag{5a}$$

$$T^1_{m\pm} = T^2_{m\pm} = T^3_{m\pm}$$

$$= \frac{4}{d}\left[(1 - \alpha_p)\alpha_p - \frac{\alpha_c^2\sqrt{V_2^2 + V_3^2}}{\epsilon^1_b - \epsilon^o_b}\right]\epsilon^1_{m\pm}, \tag{5b}$$

$$T^o_m = \frac{4}{d}\left[(1 - \alpha_p)\alpha_p - \frac{\alpha_c^2\sqrt{V_2^2 - V_3^2}}{\epsilon^o_b - \epsilon^1_a}\right]\epsilon^o_m, \tag{5c}$$

$$T^1_m = T^2_m = T^3_m$$

$$= -\frac{4}{d}\left[(1 + \alpha_p)\alpha_p - \frac{\alpha_c^2\sqrt{V_2^2 + V_3^2}}{\epsilon^1_b - \epsilon^o_a}\right]\epsilon^1_m. \tag{5d}$$

Finally combining Eqs. (1)–(5), the bond energy is

$$E_b = -1(1 + \alpha_p^2)(V_2^2 + V_3^2)^{1/2}$$

$$-\frac{9}{8}(1 - \alpha_p^2)^2\frac{(V_{1+}^2 + V_{1-}^2)}{(V_2^2 + V_3^2)^{1/2}}$$

$$+ |V_{1+} + V_{1-}| + \frac{\epsilon_{p+} - \epsilon_p}{2}. \tag{6}$$

Equation (6) was first derived by Harrison using a much simpler argument.[1] The bond energies quoted in Table I and on Fig. 5 are calculated from this expression, using as input the correlated term values from Table I and the experimental bond lengths.

B. Alloy calculations

There are several layers of sophistication that can be used to compute the bond energy and length changes between pure compounds and their alloys. However, all the calculations are motivated by the EXAFS experimental result found first for $Ga_{1-x}In_xAs$ and later in other compounds.[5] It was found that, while the average bond length in the alloy follows Vegard's rule and varies linearly between those of the pure constituents, the individual bond lengths of GaAs and

InAs are nearly the same as those of the pure materials—changing by only about 4%. The shorter GaAs bond increases its length when it is an impurity in an InAs host by about 4%, and the longer InAs bond shrinks by about 4% in the opposite extreme where it is the impurity in GaAs. The lattice fits together by having the metal sublattice nearly retain its fcc structure with a lattice constant following Vegard's rule, and each As atom adjusts its local position and accommodates to which metal atoms occupy its four neighboring sites. Thus there are five local arrangements; in the first an As atom is surrounded by four Ga atoms [*a* Ga(4)In(*o*) configuration], in the second there are three Ga atoms and one In [*a* Ga(3)In(1) configuration], and the last three configurations are Ga(2)In(2), Ga(1)In(3), and Ga(*o*)In(4). In a general cation $A_{1-x}B_xC$ alloy or anion $CA_{1-x}B_x$ alloy, the generalizations are obvious, and one can examine $A(4-\eta)B(\eta);\eta = 1,2,3,4$ configurations.

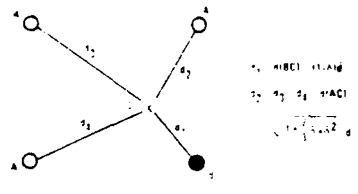
1. Method 1

The lowest level of approximation that takes account of these findings calculates [from Eqs. (1) and (2)] a given bond's energy and length shift when its surrounding bonds have different constituents. The calculation is done assuming that the surrounding bonds retain their pure crystal bond lengths. This model is somewhat unrealistic, in that it ignores the strains that must be present to allow the lattice to fit together. However, it is nonetheless useful because it treats the largest terms, viz. those arising from charge transfer between the substituted species in the alloy, and results in analytic expressions, while the successively more complicated treatments must be done numerically. It is also useful later because, taken together with the more complete calculations, it allows one to judge the relative contribution from different physical effects. From Eq. (1) in a cation-substituted alloy one can deduce an expression for the energy shift of an AC bond when η surrounding bonds are replaced by BC, ΔE_a^η as

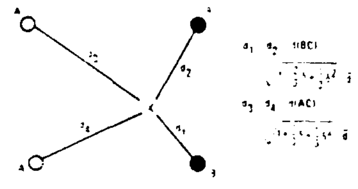
$$\begin{aligned}\Delta E_b^1 = {} & \frac{2[U^0_{b-}U^1_{a-}(\mathrm{BC})V_{1-}]^2}{\epsilon_b^0 - \epsilon_a^1(\mathrm{BC})} - \frac{2[U^0_{b-}U^1_{a-}(\mathrm{AC})V_{1-}]^2}{\epsilon_b^0 - \epsilon_a^1(\mathrm{AC})} \\ & + \frac{d\partial}{4\partial d}2\left\{\frac{[U^0_{b-}U^1_{a-}(\mathrm{BC})V_{1-}]^2}{\epsilon_b^0 - \epsilon_a^1(\mathrm{BC})} - \frac{[U^0_{b-}U^1_{a-}(\mathrm{AC})V_{1-}]^2}{\epsilon_b^0 - \epsilon_a^1(\mathrm{AC})} \right. \\ & \left. + \frac{[U^1_{b-}(\mathrm{BC})U^0_{a-}V_{1-}]^2}{\epsilon_b^1(\mathrm{BC}) - \epsilon_a^0} - \frac{[U^1_{b-}(\mathrm{AC})U^0_{a-}V_{1-}]^2}{\epsilon_b^1(\mathrm{AC}) - \epsilon_a^0}\right\}, \end{aligned} \tag{7}$$

where the partial derivative is taken with respect to the central bond length (a term designated with a superscript *o*) and in this simple approximation

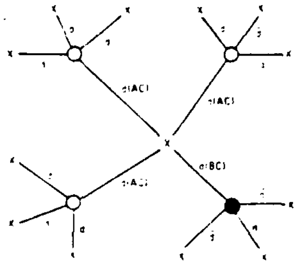
$$\Delta E_a^\eta = \eta E_b^1, \quad \eta = 0,1,2,3. \tag{8}$$



(a) A(3)B(1) CONFIGURATION



(b) A(2)B(2) CONFIGURATION



(c) A(3)B(1) EMBEDDED IN A VIRTUAL CRYSTAL

FIG. 1. Schematic view of various bond configurations. (a) A(3)B(1) configuration, (b) A(2)B(2) configuration, and (c) A(3)B(1) configuration embedded in a virtual crystal.

The bond length shifts are found from Eq. (2). The modification of these expressions for an anion-substituted alloy follows from symmetry.

2. Method 2

A more realistic calculation includes strain and permits the lattice to fit together. The simplest non-self-consistent, near-neighbor version of this calculation proceeds as follows. In a cation-substituted alloy $A_{1-x}B_xC$ (e.g., $Hg_{1-x}Cd_xTe$ or $Zn_{1-x}Hg_xTe$), the cations are assumed to occupy their regular fcc sublattice sites, while the anions accommodate to their local configuration. The average lattice constant is taken to follow Vegard's rule:

$$\bar{d} = (1-x)d_0(\mathrm{AC}) + xd_0(\mathrm{BC}), \tag{9}$$

where d_0 and $d_0(\mathrm{BC})$ are the pure crystal lattice constants. Then in an A(3)B(1) configuration [see Fig. 1(a)], the anion will be displaced along the BC bond. For definiteness, suppose $d(\mathrm{AC}) > d(\mathrm{BC})$; then the anion will be shifted away from the tetrahedral site toward the B atom. Because $d_0(\mathrm{AC}) > \bar{d} > d_0(\mathrm{BC})$, the $d(\mathrm{BC})$ bond will be elongated and the $d(\mathrm{AC})$ bonds will also be locally elongated in a strained configuration. We can for this case [see Fig. 1(a)] write

$$d(\mathrm{BC}) = (1-\delta)\bar{d} \tag{10}$$

and for the three AC bonds

$$d(\mathrm{AC}) = \sqrt{1 + \frac{2\delta}{3} + \delta^2}\,\bar{d}. \tag{11}$$

In an A(2)B(2) configuration [see Fig. 1(b)], we have for both BC bonds

$$d(\mathrm{BC}) = \sqrt{1 - \frac{2}{3}\delta + \frac{1}{3}\delta^2}\,\bar{d}, \qquad (12)$$

and for both AC bonds

$$d(\mathrm{AC}) = \sqrt{1 + \frac{2}{3}\delta + \frac{1}{3}\delta^2}\,\bar{d}. \qquad (13)$$

The values of δ are determined by the condition $\partial E_T/\partial\delta = 0$, or

$$\sum_{n=0}^{4}\left(T_{on} + T_{mn} - \frac{4\gamma_n}{d_n^5}\right)\frac{\partial d_n}{\partial\delta} = 0. \qquad (14)$$

Equation (14) is solved by

(1) Assigning the γ_ns their pure crystal values.
(2) Supposing that the four tetrahedrally configured anions are embedded in the virtual crystal [see Fig. 1(c)].
(3) For each x, δ is varied and the bond tensions are calculated and inserted into Eq. (14) until a value of δ is found that satisfies this equation.

The computer outputs are δ and the AC and BC bond energies. This method can be extended to take explicit account of the bond length variations of larger clusters, and $\bar{d}$ can be calculated explicitly (instead of using Vegard's rule) so the results can be made to be self-consistent. Also, effects arising from bond angle distortions can be included. The major effects that are currently included in the calculation arise from the bond length dependence of V_2, and charge shifts driven by hybrid energy differences and coupled through the metallization terms. However, V_2 is also sensitive to bond angle distortions, and these effects are not included in the present results. We expect their inclusion will modify the quantitative results by 10%–30%, but not the trends.

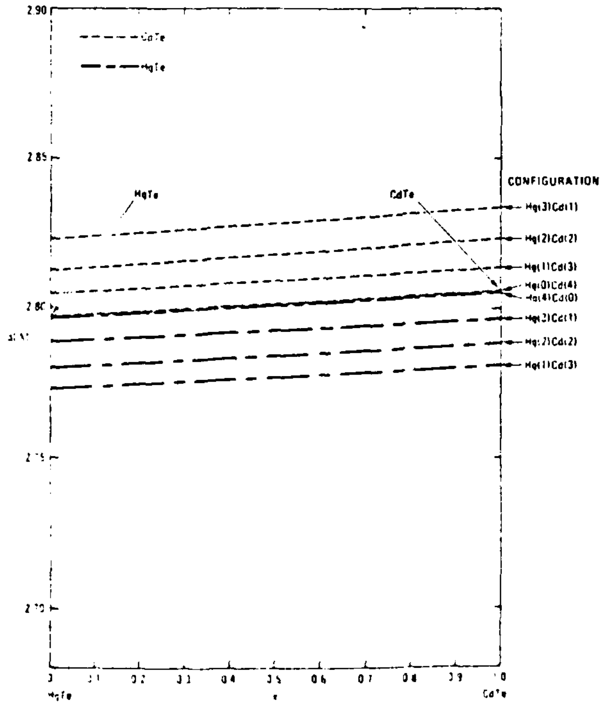


FIG. 2. Bond lengths as a function of concentration x for the $\mathrm{Hg}_{1-x}\mathrm{Cd}_x\mathrm{Te}$ alloy in different configurations. The short dashed lines are for the CdTe bonds and the long–short dashed lines represent the behavior of the HgTe bonds.

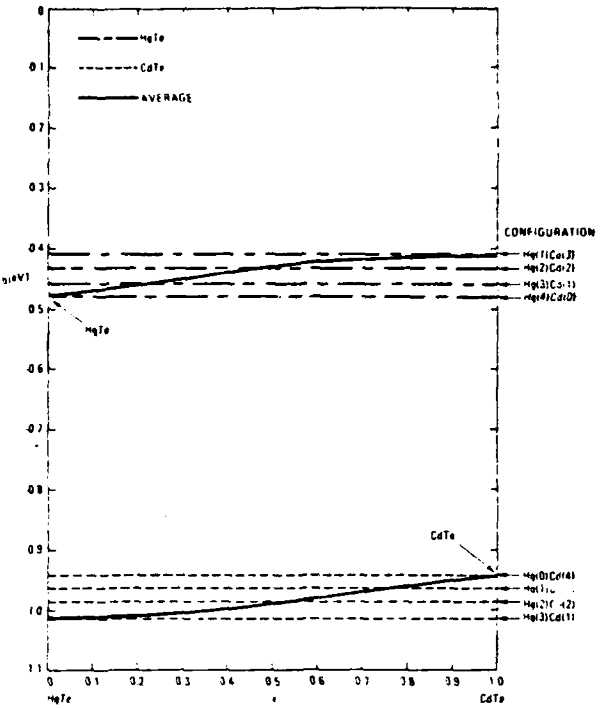


FIG. 3. Bond energy as a function of concentration x for the $\mathrm{Hg}_{1-x}\mathrm{Cd}_x\mathrm{Te}$ alloy in different configurations. The short dashed lines are for the CdTe bonds and the long–short dashed lines represent the behavior of the HgTe bonds. The solid lines are the concentration weighted averages of the respective bond energies.

C. Results

The correlated atomic term values[4] for the elements in the more common II–VI compounds are listed in Table I, along with the bond energies of their nine compounds. The valence s state energies for Zn are large, become smaller as for Cd, and then (in an unusual occurrence) the trend reverses and the Hg s levels are deeper again. This is caused by the relativistic terms that become important in Hg. The p state levels have a similar trend; however, the differences among the atoms are much smaller. The levels of the anions are deeper than those of the cations. Thus, there is a net electron transfer from the cations to the anions, which is responsible for the ionic contribution to the bonding. The polarity, also listed in Table I, reflects the relative contribution of ionic and covalent character to the bonds. Notice that α_p is small for ZnTe larger for CdTe, and smaller again for HgTe. By contrast, the metallic contribution E_m is large for ZnTe and HgTe and small for CdTe. In fact, HgTe would not be bound if it were not for the metallic contribution to the bond, not a terribly surprising result, because HgTe is a semimetal. The bond lengths of CdTe and HgTe are very nearly the same, but this is an accident resulting from the balance of contributions to their bonds: CdTe is dominated by the covalent and ionic contributions to its bond, while HgTe has bonds that are mostly covalent and metallic. We should therefore not be too surprised if their alloy displays some unusual bond length variations.

Figures 2 and 3 illustrate the $Hg_{1-x}Cd_xTe$ alloy concentration variations of the length and energy for the five configurations. These results were calculated using the second method discussed in Sec. II B. In Fig. 2 we see the unexpected result that the shorter HgTe bond becomes still shorter in the alloy, while the CdTe bond length increases. Moreover, the changes are large compared to the original length difference. Thus, one can expect to find local microstrains, even in this nearly lattice-matched material. The bond energy variations shown in Fig. 3 also display an important result. The already weak HgTe bond is destabilized by the presence of Cd. The flat concentration variation of the bond lengths and energies is peculiar to HgCdTe; much more structured behaviors are exhibited by other compounds.

Several considerations to keep in mind when dealing with alloys are depicted schematically in Fig. 4 for a Hg(3)Cd(1) configuration. Due to the charge shift alone, the HgTe bond shrinks by 0.030 Å, the minimum in the configuration diagram moves to a smaller bond length and the depth of the well is decreased. Then the bond is stretched by 0.011 Å to its final length. Thus, the bond is not centered at a local minimum. For example, this implies that the force constant of local phonon modes will be related to the second derivative of this configuration diagram evaluated at the strained position away from the well minimum. This effect is even more pronounced for materials with a larger lattice mismatch. Notice that the CdTe bond is also stretched relative to its local minimum in this configuration.

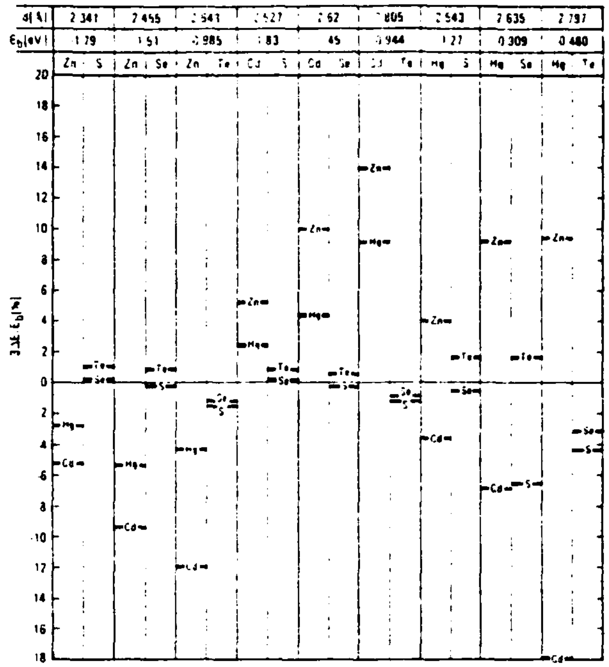


FIG. 5. AC bond energy shifts for the II–VI compounds in A(1)B(3) configurations expressed as a percentage of the pure crystal bond energy. The B atom is designated in the line that shows the shift. The pure crystal AC bond energies and lengths are given at the top of the figure. These numbers were calculated by method 1 in Sec. II B.

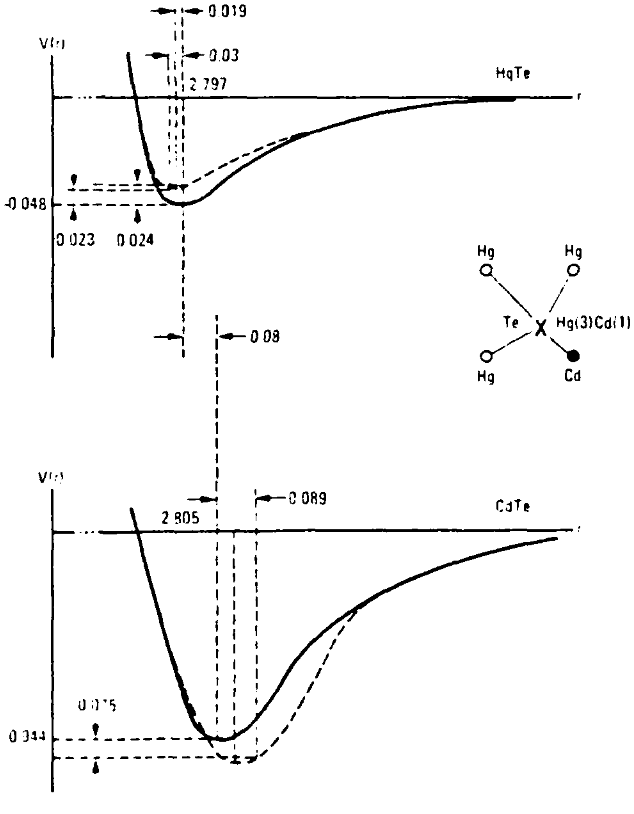


FIG. 4. Schematic picture of a potential configuration diagram for the HgTe and CdTe bonds in an Hg(3)Cd(1) configuration. For the HgTe bond the energy and position of the bond center are calculated by method 1 of Sec. II B and the final stretched position and energy by method 2. The results quoted for the CdTe bond are those determined by method 2.

The result for bond energy variations for all the II–VI compound alloys is gathered in Fig. 5. Here the energy shifts of an AC bond in a A(1)B(3) configuration $\Delta E_b^3 = 3\Delta E_b^1$ are presented as a percentage of the pure material bond energy for cation substitute alloys along with equivalent results for anion substantial alloys, e.g., $HgS_{1-x}Te_x$. The energy shifts are calculated following the first method presented in Sec. II B. We have not had the opportunity to run all the compounds following the more sophisticated calculational procedure. The bond at the top of a column is the one whose energy shift is calculated, and the other constituent is identified in the line that indicates the energy shift. A negative $3\Delta E_b/E_b$ is destabilizing and a positive one stabilizing. Focus on HgTe. It is destabilized by CdTe, HgS, and HgSe, but is stabilized by ZnTe. The CdTe destabilizes the HgTe bond by 18% according to the simple calculation and by 15% according to the better one. The difference between Zn and Cd occurs because in the Cd case there is a net electron transfer from the Cd to the Hg, while in the case of Zn there is practically no such transfer. These extra transferred electrons must occupy antibonding states, because the bonding states on the HgTe are full, so the net effect (including other energy state shifts) is to destabilize the HgTe bond. This predicted destabilization of the HgTe bond in HgCdTe alloys

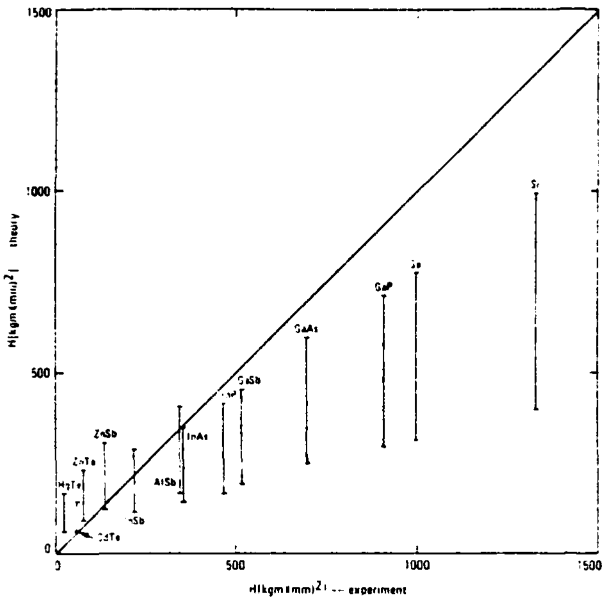


FIG. 6. Theoretical vs experimental hardness. The two values at the top and bottom of the vertical lines designated for each compound were calculated for two different possible dislocation arrangements. The lower one corresponds to the minimum energy configuration and the upper one to a likely higher energy configuration. The experimental numbers were taken from Ref. 2.

agrees with the *d* state energy shift reported by Spicer *et al.* at last year's workshop,[6] and by the oxygen uptake experiments presented by the same group at this conference.[7]

III. DISLOCATION ENERGIES AND HARDNESS

The dislocation energies and hardness of semiconductor are calculated by an extension of Harrison's method.[1] It is demonstrated, in agreement with experiment, that dislocation energies per unit length, depending on the covalency, are proportional to d^{-9}, where d is the bond length and hardness is proportional to d^{-5} to d^{-11}. The low powers are for pure covalent materials. The hardness is related to the interaction energies among dislocations. The detailed theory will be published elsewhere.[8]

Figure 6 shows the theoretical hardness of several semiconductors plotted against experimental values.[2] The theory is calculated for two different dislocation configurations caused by the extra material pushed by the indentor into the crystal. The lower end of each vertical line is a realistic hardness for the minimum energy configuration, and the upper ends of the lines are the values found for a higher-energy configuration. In most cases, one would expect real configurations to have hardnesses lying somewhere between these two values, once the effect of heat generated by the indention process is subtracted from the experimental numbers. This theory has no adjustable parameters, yet it predicts the right trends and magnitudes.

IV. NATIVE IMPERFECTIONS GROWN INTO ALLOYS

There are two kinds of imperfections we shall discuss: vacancies and dislocations. At present, the conclusions are tentative, because the detailed theory is still incomplete. Vacancy formations energies and the melting temperature are both proportional to the bond energies of the constituents. Thus, one would expect the vacancy density just below the melting point to be about the same for all materials with a similar coordination number: in our semiconductor case, the coordination number is four. Dislocation energies are mostly owing to long-range bond angle distortions[9] and accordingly are sensitive to the sheer coefficients, which vary roughly as d^{-7} for many compounds.[1] Based on this general argument, one would expect to find fewer dislocations in materials with short bond lengths.

Dislocations appear in bulk grown materials through a sequence of steps.[10] We shall only discuss the simplest case, where there are no mechanical stresses on the growing material. In the temperature gradient behind the growth front, there is usually a supersaturated vacancy content. These excess vacancies may diffuse to surfaces where they do no harm, or they may condense into vacancy clusters that can then organize into dislocation loops. If a loop is not parallel to the growth front, it can climb toward the front by absorbing more vacancies. If it climbs faster than the growth progresses, then the loop will reach the front. It can then spread and grow along with the crystal. Observed from the growth surface, such a loop will appear to be two edge dislocations. If this dislocation formation process can be inhibited in any of its steps, the resulting material will have fewer dislocations.

Zn in $Zn_{1-x}Cd_xTe$ or $Zn_{1-x}Hg_xTe$ will serve this function in two ways. The first is that Zn stabilizes the weaker bond in these alloys; accordingly, vacancy formation energies in the alloy should be larger than those in the pure constituent CdTe or HgTe. The melting temperature will also be raised somewhat; which effect "wins" is a matter still to be decided. The second effect is clear cut: The shorter bond lengths of the ZnTe will introduce stiff struts into the system that will increase the dislocation energies and thus should inhibit dislocation loop formation and subsequent climb. The large dislocation density reduction found by Bell and Sen in $Zn_{1-x}Cd_xTe$ with only 4% Zn supports this speculation.[11]

V. CONCLUSIONS

We have argued that, in an alloy, the bond energies and bond lengths are modified from their pure crystal values. Bond energies can be strengthened or weakened in an alloy depending on the relative sp^3 state energies and the consequent net electron shift between the constituents. For most compounds, the average bond length modifications for each type of bond are generally small compared to those suggested by Vegard's rule, in agreement with experiment. However, HgCdTe is an exceptional case, because the bond lengths for HgTe and CdTe are nearly the same (by accident). The bond length and energy shifts also depend sensitively on the local configuration and the bonds are locally strained. These effects must be taken into consideration in a proper theory of alloy vacancy formation energies, phonon frequencies, dislocation energies, etc.

The conclusions most relevant to the initial question posed in the introduction are that for the narrow gap allows:

Cd destabilizes the already weak HgTe bond and has little effect on the dislocation energy.

Zn both stabilizes the HgTe bond and increases the dislocation energy.

Zn also decreases the dislocation density of CdTe to improve substrates.

HgZnTe electron and hole mobilities are about the same as those of comparable band gap HgCdTe.

HgZnTe is likely to be a superior material to HgCdTe for detectors operating in the 8–12 μm regime.

---

[a] This work was supported in part by AFOSR Contract No. F496280-81-K-0012 and by DARPA Contract No. MDA-903-83-0-0108.

[b] Stanford Ascherman Professor of Engineering.

[1] W. A. Harrison, *Electronic Structure and the Properties of Solids* (Freeman, San Francisco, 1980); R. C. Sokel, thesis, Stanford University, 1978; W. A. Harrison, *Microscience* (SRI International, Menlo Park, California, 1983), Vol. IV, p. 34 limited-distribution.

[2] N. A. Goyunova, A. S. Borshchevskii, and D. N. Fretaikov, *Semiconductors and Semimetals*, edited by R. L. Willardson and A. C. Beer (Academic, New York, 1958), Vol. IV, Chap. I.

[3] O. Mukhina, A. M. Gulyaev, A.S. Shnitnikov, and I. V. Yamshchikova, Tr. Mosk Energ. Inst. **443**, 16 (1980).

[4] A. B. Chen, A. Sher, and W. E. Spicer, J. Vac. Sci. Technol. A **1**, 1674 (1983).

[5] J. C. Mikkelsen, Jr. and J. B. Boyce, Phys. Rev. Lett. **19**, 1412 (1982), and private communications.

[6] W. E. Spicer, J. A. Silberman, I. Landau, A. B. Chen, A. Sher, and J. A. Wilson, J. Vac. Sci. Technol. A **1**, 1735 (1983).

[7] J. A. Silberman, D. Laser, I. Landau, and W. E. Spicer (these proceedings).

[8] A. Sher, A. B. Chen, and W. E. Spicer, Appl. Phys. Lett. **46**, 54 (1985).

[9] J. P. Hirth and J. Lothe, *Theory of Dislocations*, 2nd ed. (Wiley, New York, 1982).

[10] G. Schoeck and W. A. Tiller, Philos. Mag. **5**, 43 (1960).

[11] S. L. Bell and M. Sen (these proceedings).

[12] R. Zallen, *Handbook of Semiconductors*, edited by T. S. Moss and W. Paul (North-Holland, Amsterdam, 1982), Vol. I, Chap. 1.

PHENOMENA INFLUENCING THE DISLOCATION DENSITY OF SEMICONDUCTOR COMPOUNDS AND ALLOYS[a]

A. Sher, SRI International, Menlo Park, CA 94062
An-Ban Chen, Auburn University, Auburn, AL 36489
W.E. Spicer[b], Stanford University, Stanford, CA 94305

Abstract

The objective of this paper is to identify the principal microscopic phenomena controlling dislocation densities in bulk grown semiconductors. Then, based on this understanding, a strategy for selecting materials to reduce dislocation densities is offered. The relevant quantities are calculated from an extension of Harrison's bonding theory, which, with our improved accuracy relates properties of the solids to the constituent atoms' valence electron energy states and wave functions. We report on the alloy composition variation of bond energies, bond lengths, charge redistribution among constituents, vacancy formation energies, dislocation energies, and hardness. Several III-V and II-VI compound semiconductors are treated including, GaAs, GaInAs, HgCdTe, and ZnHgTe.

(a) This work was supported in part by AFOSR Contract F49620-81-K0012 and DARPA Contract MDA-903-83-C-0108.

(b) Stanford Ascherman Professor of Engineering

Introduction

According to a currently accepted model (1) of the mechanism leading to dislocations in bulk-grown material at a given temperature and temperature gradient behind the growth front; their density is governed by:

- The vacancy density that depends on the *formation energies*,
- The competition between vacancy annealing rates and vacancy interaction caused clusters,
- The condensation rate of these vacancy clusters into dislocation loops, and
- The subsequent growth rate of these loops.

The objective of this paper is to identify some of the principal microscopic phenomena controlling these features in semiconductors compounds and their pseudobinary alloys. If any of the foregoing steps can be inhibited, then there will be fewer dislocations in bulk-grown crystals.

A model of the bonding of tetrahedrally coordinated semiconductor compounds due to Harrison underlies this work (2). He has derived expressions for the bond energies and strain coefficients of pure semiconductor compounds in terms of the constituent atom's valence state wave functions and energies. There are four contributions to the bond energies:

- A covalent energy arising from the interaction between sp^3 hybrids on adjacent sites, which, according to a universal rule deduced by Harrison, is related to the inverse square of the bond length, d,
- An ionic energy which is proportional to the energy difference between the sp^3 hybrid energies of the anion and cation,
- A metallic energy arising from the interaction between filled bonding and unfilled antibonding states on adjacent bonds, and
- An overlap repulsion energy which is taken to vary as d^{-4}.

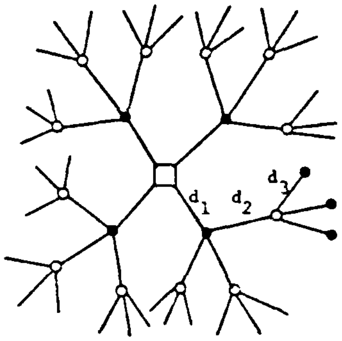
The shear coefficients, which play an important role in dislocation energies, are shown to vary for covalent solids as the covalent energy per cell unit volume or as d^{-5}. The ionic energy arising as it does from coulomb interactions, is insensitive to bond angle distortions and depends only on bond lengths. Thus, in the ionic semiconductors the shear coefficients are reduced from those of equivalent bond length covalent compounds, but also the power law dependence on bond length increases to d^{-11} for the extreme ionic limit (3). The metallic interaction is bond angle sensitive and contributes to the d^{-11} power law.

We have modified Harrison's theory in several ways (3) (4). First, instead of using atomic term values as the input energies in the bonding calculation, we use correlated atomic energies that take account of all the atomic electronic energy changes associated with a state change. In our prior publications these energies were calculated from pseudopotentials (3). In this paper we have modified the procedure and taken the ground state energies from free atom ionization experiments, and only calculate the excitation energies from the pseudopotentials. Second, the Harrison's theory has been extended to alloys in which each type of the individual bond's energy and length changes and the net alloy substitution energies are calculated. A theory of dislocation energies and hardness of semiconductors has been advanced.(4) In this paper the second item will be emphasized since item one is less relevant to this conference and these will be published elsewhere.

Bond Length and Energy Changes in an Alloy

Single Impurity

Start by considering the simplest defect, one cation of a host semiconductor compound AC is replaced by an isoelectronic impurity I, as sketched in Fig. 1. This will be generalized to an $AB_{1-x}C_x$ alloy later ans the cation substituted case $A_{1-x}B_xC$ follows by symmetry. To study the principal effect of core lattice distortions, we allow the positions of the first shell of atoms labeled by bond length d_1 to move, but fix the atoms of the second neighbors and beyond at their pure-crystal positions. Eventually, the effects of long range strain fields, generated by a point distortion, must be added to this core distortion. Here we allow only three different bond lengths $d_1 = (1-\delta)d$, $d_2 = (1 + 2/3\delta + \delta^2)^{1/2}d$, and $d_3 = d$, where d is the bond length of the pure host AC compound, and δ is a scale parameter.



Schematic representation of the bonds around an atom of interest (designated by the square). The first, second, and third neighbor bonds are designated d_1, d_2, and d_3

There are four bonds with energy E_1, twelve with energy E_2, and 36 with energy E_3. These energies can be written in closed form in the Harrison theory. The energies E_1 and E_2 differ from the host bond energy E_o because of bond length and angle distortions as well as difference in the ionic energy of I and C atom induced charge transfer. The energy E_3 differs from E_o only by the charge transfer coupled through the metallization terms. There are two interesting energies we can calculate. The first is the energy Δ_s required to substitute the I atom for a C atom, i.e. bring a free I atom from infinity and replace a C atom that is taken from the crystal and removed to infinity.

$$\Delta_s \equiv 4\Delta E_1 + 12\Delta E_2 + 36\Delta E_3 - (\epsilon_C - \epsilon_I) \qquad (1)$$

where $\Delta E_j = E_j - E_o$, $j = 1,2,3$, and ϵ_C and ϵ_I are the free atom valence electron binding energies. The second is the bond energy change of the impurity AI bond relative to the bond energy of a pure AI compound, denoted Δ_I

$$\Delta_I \equiv \Delta E_1 - \frac{1}{4}(\epsilon_I - \epsilon_C) \qquad (2)$$

This energy tells us if the AI bond is stabilized ($\Delta_I < 0$) or destabilized ($\Delta_I > 0$) when it is in an AC host. If we define Δ_b as the bond energy difference between the IA and CA bonds each in their respective pure crystals

$$\Delta_b \equiv BE(IA) - BE(CA) \qquad (3)$$

then one can write

$$\Delta_s = 4\Delta_b + \Delta_{sc} \qquad (4)$$

where Δ_{sc} is extra energy difference caused by strain and charge transfer.

The energyes E_j and bond lengths d_j, $j = 1, 2, 3$, are calculated by minimizing Δ_s. If we let $d_I = d_o\,(1 - \delta_o)$ where d_I is the bond length in the pure AI lattice, then the approximate expression

$$\delta_I \approx \frac{\delta_o}{1 + \frac{1}{3}\frac{B}{B_I} + \frac{8(C_{11} - C_{12})}{27\,B_I} + \text{(small terms due to charge transfer)}} \qquad (5)$$

can be deduced. The bulk modulii B of the host and B_I of the impurity lattices, and a shear coefficient $C_{11} - C_{12}$ of the host appear in the expression. While this expression is approximate, the effects of the various types of strains can be visualized using it. Our detailed calculations are done using the full theory. If one fixes the bond lengths $d_1 = d_2 = d_3 = d_o$, then $\delta = \delta_o$ and the strain energy Δ_{sc} is large and positive. If one allows only bond length distortions ($C_{11} - C_{12} = 0$) then δ is reduced and the bonds tend to adjust so $d_1 \approx d_I$. This reduces Δ_{sc} considerably from the undistorted lattice case. If $d_I > d_o$ then both bonds are stretched somewhat, $d_1 - d_I > 0$ and $d_2 - d_o > 0$. However, this configuration produces large bond angle distortions. When the shear coefficients are turned on the lattice relaxation is modified and the bond lengths cannot adjust as much, so d_1 differs from d_I by a larger amount. The net result is that Δ_{sc} is increased since one pays the price of strain energy either as a bond length or bond angle distortion.

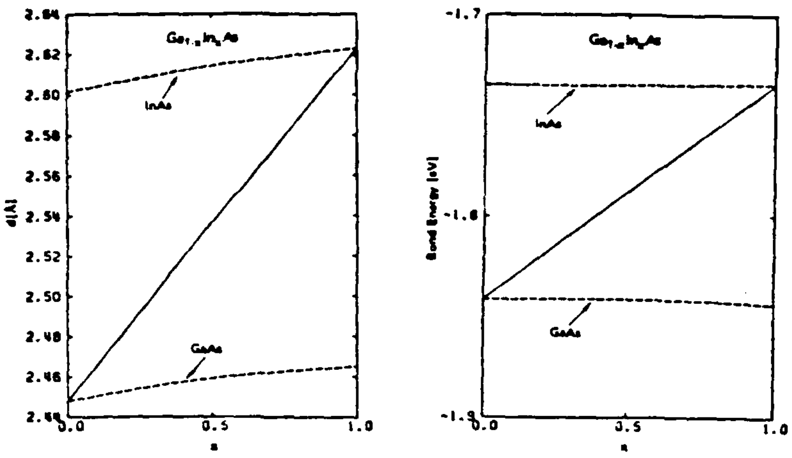
If we calculate δ from the full theory, then we predict values that are too large. This occurs because the theory predicts $C_{11} - C_{12}$ which agree well with experiment but it predicts bulk modulii with the proper trends from one compound to another but which are about a factor of 2.4 too small. If we use the experimental values of the strain coefficients in Eq. (5) then good agreement is found with the experiments (5) on $Ga_{1-x}In_xAs$ and $ZnSe_{1-x}Te_x$. The results are quoted in Table I.

Table I. Bond Lengths in Å for Impurities in the Indicated Host

	Ga in InAs	In in GaAs	Se in ZnTe	Te in ZnSe
Experiment$^{(5)}$	2.487	2.587	2.496	2.595
Eq.(5) and experimental B&C	2.499	2.547	2.510	2.570
Full Theory	2.538	2.518	2.570	2.512

Alloy

In an $A_{1-x}B_xC$ alloy the four cations around a given C anion can be arranged in five different configurations, denoted by $A(4-\eta)$, $B(\eta)$, $\eta = 0,1,2,3,4$. An A(2) B(2) configuration, for example, is one in which the C atom has two A atom and two B atom neighbors. It is possible to solve the full alloy problem for large clusters, but for now we have restricted the cluster to near neighbors only, and in Fig. 1 the central atom is now an anion C and $d_2 = d_3 - d_{eff}$ are taken to be an effective medium bond length that is determined self-consistently. The different bonds of type 1 no longer have the same length. The procedure is as follows: First assume a value for d_{eff}, say the virtual crystal bond length $(1-x)\,d\,(AC) + xd\,(BC)$. Next calculate the d_i values for the various configurations by minimizing the energy of the configuration. Then, configuration average the various d values to find a new d_{eff}. Then iterate this procedure until it converges.



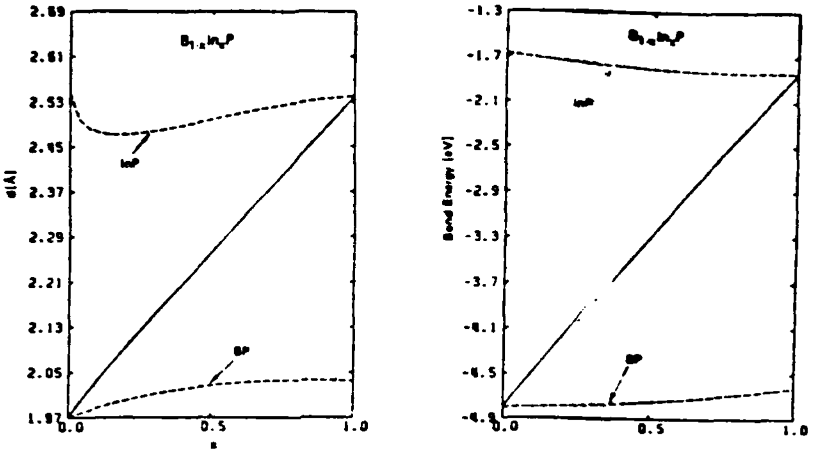
(a) (b)

Figure 2: Bond length (a) and energy (b) of $Ga_{1-x}In_xAs$ as functions of x. The dashed curves are the average values for the designated bonds, and the solid curved are the alloy averages.

Because the bond angle distortion terms are unphysically large we have left them out of the present calculation. Hence bond length shifts will be somewhat too small.

Results for the bond energy and bond length shifts with concentration are given for $Ga_{1-x}In_xAs$ in Fig. 2 and $B_{1-x}In_xP$ in Fig. 3 as examples of two different behavior patterns. The predicted trends for GaInAs agree with experiment but the bond length changes are too small as expected with the bond angle distortion terms absent. Notice that the longer bond in this case decreases in length and the shorter one increases as expected. However, in the BInP case the charge shift terms are so large that the longer InP bond has a minimum as a fucntion of composition rather than a monotonic decrease. Because of the large bond length difference between BP and InP there is probably a missability gap in this alloy that prevents these compositions from being prepared. However, one may be able to examine the variation of the anomalous behavior of the InP bond in a BP host (x << 1).

The $Hg_{1-x}Cd_xTe$ system is completely anamolous. The bond lengths of CdTe (2.805 Å) and HgTe (2.797 Å) are nearly the same by an accident. CdTe bonds are dominated by covalent and ionic interactions, while HgTe is more covalent and the metallic terms are important. The mix of interactions in the two cases leads to the same bond lengths. When an alloy is made the charge shift terms dominate, and they cause the longer CdTe bond to become still longer and the HgTe to contract by amounts that are large compared to the pure crystal difference. Moreover, the already weak HgTe has its bond energy reduced still more by the presence of Cd. Since the melting point of HgTe increases as Cd is added and the local strength of HgTe bonds adjacent to a Cd decrease, vacancy densities will increase. All these and other observed phenomena are predicted by the theory. A complete catalog of results requires more pages than we have been allocated in this article, but we have tried to offer a representative group.



(a) (b)

Figure 3: Bond length (a) and energy (b) of $B_{1-x}In_xP$ as functions of x. The dashed curves are the average values for the designated bonds and the solid curves are the alloy averages.

Conclusions

We have demonstrated in agreement with experiment that dislocation energies and hardness of semiconductors are proportional to the shear coefficients which vary roughly as d^{-9}. Thus to decrease dislocations in a given semiconductor one should inhibit their formation by introducing some means of shortening bonds. This should be accomplished without also decreasing the ratio of the vacancy formation energy to the melt temperature. While detailed calculations to support the following contentions are still incomplete, a set of criteria on an impurity (denoted I) in a host semiconductor (denoted H) that are likely to satisfy these conditions are: the bond length of the impurity is smaller than that of the host $d_I < d_H$, the smaller bond energy is stabilized in the alloy $\min\{\Delta E_{bH}, \Delta E_{bI}\} < 0$. In these circumstances, the average bond length will shrink and in the vicinity of each impurity the four surrounding bonds nearly have the length of the impurity but are stretched slightly. The next neighboring host bonds are also stretched. The net effect is an arrangement which is more rigid than the unperturbed lattice and consequently the local shear strain energy increases, causing dislocation energies to increase. Examples of this case are $B_{1-x}Ga_xAs$, and $Zn_{1-x}Hg_xTe$. Both B in GaAs (6) and Zn in HgTe (7) have proven to be effective in reducing dislocation densities. Unfortunately, B has a low solubility in GaAs and it is not clear that enough can be gotten in to make it easy to prepare dislocation-free material.

A second case where an improvement occurs is if $d_I > d_H$, and again $\min\{\Delta E_{bH}, \Delta E_{bI}\} > 0$. In this case each of the four longer impurity atom bonds are compressed by the surrounding host bonds and, more importantly, the twelve next neighbor host bonds are also compressed. Once again for small impurity concentrations the net dislocation energy should be increased. However, in this case the effect is competing against a net bond lengthening trend of the alloy which is tending to make it less rigid against a shear, so at higher concentrations the mechanism should cease to function. An example of this case is $Ga_{1-x}In_xAs$, where In has proven to be effective in dislocation reduction of GaAs (8).

We have demonstrated that our modification of the Harrison bonding theory accruately predicts the observed change of bond lengths in semiconductor alloys and offers guidance to means for dislocation reduction.

References

1. G. Schoeck and W.A. Tiller, Phil. Mag. 5, 43 (1960).
2. W.A. Harrison, *Electronic Structure and Properties of Solids*, (Freeman, San Francisco, 1980); R.C. Sokel Thesis, Stanford University, 1978; W.A. Harrison Microscience , Vol. 4 (limited distribution SRI International publication 1983) p. 34.
3. A. Sher, An-Ban Chen, and W.E. Spicer "Effects Influencing the Structure Integrity of Semiconductors and Their Alloys", 1984 Workshop on the Physics and Chemistry of Mercury Cadmium Telluride, San Diego, 1984.
4. A. Sher, An-Ban Chen, and W.E. Spicer, "Dislocations Energies and Hardness of Semiconductors", (submitted to Appl. Phys.Lett.).
5. J.C. Mikkelsen, Jr. and J.B. Boyce, Phys. Rev. Lett. 49, 1412 (1982), and private communications.
6. G. Jacob, J. Crystal Growth 59, 669 (1982); Y. Seki, H. Watanabe, and J. Matsui, J. Appl. Phys. 42, 822 (1983).
7. S.L. Bell and S. Sen, IRIS Detector Specialty Group Meeting, Boulder Colorado (1983); T.W. James and B.F. Zuck, ibid.
8. R.N. Thomas, H.M. Hobgood, D.L. Barrett, G.W. Eldridge, M.M. Sopira, and M.C. Driver, "Large Diameter, Low-Dislocation In Doped GaAs:..." Third Conference on Semi-Insulating III-V Materials, Warm Springs, Oregon 1984.

SEMICONDUCTOR ALLOYS: LOCAL BOND LENGTHS, MIXING ENTHALPIES, AND MICROCLUSTERS

A.B. CHEN* AND A. SHER†
*Physics Department, Auburn University, AL 36849
†SRI International, Menlo Park, CA 94025

ABSTRACT

Several recent theoretical studies of the local structure of semiconductor alloys are summarized. First, dilute limit calculations of local bond lengths and mixing enthalpies are discussed. These calculations include effects due to both bond length and bond-angle distortions, as well as local chemical rearrangements. Then, a new statistical theory of concentrated alloys is described. Deviations from random alloy distributions (microclusters) are predicted.

INTRODUCTION

This paper summarizes our recent theoretical studies directed toward understanding the microscopic structures of pseudo-binary semiconductor alloys $A_xB_{1-x}C$. We first present a detailed calculation of the local bond length relaxation in the dilute limit $x \rightarrow 0$, i.e. the case where an A atom is substituted for a B atom in a BC compound. The mixing enthalpy parameter Ω is found to be related to small excess substitution energies. These excess energies are calculated directly through a minimization procedure. Thus, the accuracy of the predicted Ω is not limited by trying to find small differences between large numbers. The theory is then generalized to concentrated alloys using statistics based on combinations of tetrahedral clusters of five atoms. Our results predict that microclustering occurs in a majority of alloys. We conclude by identifying systematic correlations between the theory and several experiments.

Before discussing the calculation, it is useful to provide some background about the structure of these alloys. It was customary to assume that these alloys have two sublattices in which the C atoms occupy one sublattice, and A and B atoms are randomly distributed on the other. This picture, referred to as the virtual crystal approximation (VCA), implies that the nearest-neighbor (nn) bond lengths in the alloy are the concentration weighed average values, i.e. $d_{AC} = d_{BC} = \bar{d} = x\, d_{AC}^{(0)} + (1-x)\, d_{BC}^{(0)}$ where the values with a superscript (0) denote the pure-crystal values. On the other hand, according to Pauling's covalent radii approximation (CRA), the local bond lengths retain their respective pure-crystal values, i.e. $d_{AC} = d_{AC}^{(0)}$ and $d_{BC} = d_{BC}^{(0)}$

If we define $\delta_0 = (\bar{d} - d_{AC}^{(0)}) / \bar{d}$ and $\delta = (\bar{d} - d_{AC}) / \bar{d}$, then the ratio δ/δ_0 in VCA is zero, but in CRA it is 1. However, Mikkelsen and Boyce(1) found from their EXAFS experiment on $Ga_xIn_{1-x}As$ that the nn bond lengths do not fit either VCA or CRA. Instead, they found the value of δ/δ_0 to be close to 3/4. Since then, similar experiments have been done for a number of zinc-blende pseudo-binary alloys,(1) and the 3/4 rule appears to be quite general.

DILUTE LIMIT

The dilute limit is the easiest case but is still not trivial. Its solution provides both end-point results ($x = 0$ and 1), as well as insight into the extension to the concentrated alloy case. A complete description of this case is being reported elsewhere; here we summarize the essential results. The substitution energy Δ_s for an A atom replacing a B atom in a BC compound is calculated and minimized to find the relaxed configuration. Δ_s can be written as $\Delta_s = 4\,(\Delta E_b + \Delta E_s + \Delta E_{ch})$, where ΔE_b is the binding energy difference between the AC and BC compounds, ΔE_s is the strain energy, and ΔE_{ch} is a chemical energy shift. All $\Delta E'$ s are energies per bond. Then, $\Delta E = \Delta E_s + \Delta E_{ch}$ is the excess energy per bond for the impurity substitution. ΔE_s is calculated by dividing the crystal into two regions. Outside R (which is the distance of the second-shell atoms to the impurity), the distorted crystal is treated as an elastic continuum with a radial displacement field which is inversely proportional to the square of the radial distance, so $\Delta E_s^{(out)} = 1/4RCu^2$, where C is an effective shear coefficient,

$$C = \pi\,(1.6\,(C_{11} - C_{12}) + 4.8\,C_{44}), \tag{1}$$

and u is the magnitude of the displacement at R. Inside R, the strain energy $\Delta_s^{(in)}$ is treated with a valence force field (VFF).(2) Finally, the chemical energy shift ΔE_{ch} is calculated from Harrison's model and arises from changes in the metalization energies caused by different bond lengths $\Delta d \equiv d_{BC} - d_{AC}$ and covalent energies $\Delta V_3 \equiv V_3(AC) - V_3(BC)$. Note that $\delta_0 = (d_{BC}^{(0)} - d_{AC}^{(0)}) / d_{BC}^{(0)}$ and $\delta = (d_{BC}^{(0)} - d_{AC}) / d_{BC}^{(0)}$ in this dilute limit, so the excess energy ΔE can be expanded up to second order in δ, u, and ΔV_3 For a given pair A and B, ΔE is an explicit function of δ and u. Minimization of ΔE with respect to δ and U leads to the equilibrium local bond length d_{AC} and energy ΔE. Then, ΔE is used to estimate the mixing enthalpy parameter Ω in the mixing enthalpy $\Delta H = x(1-x)\Omega$ by

$$\Omega = 2\,(\Delta E\,(A \text{ in } BC) + \Delta E\,(B \text{ in } AC)). \tag{2}$$

A systematic comparison with other models based on strain energy alone shows that an increase of the range of the fixed boundary R increases the relaxation of d_{AC}, i.e. it causes δ/δ_0 to increase. The inclusion of the bond angle restoring force, on the other hand, reduces the relaxation. It turns out that a delicate cancellation of these two effects causes a simple spring model pointed out by Shih et al. (SSHS)(3) to yield accurate results. In this model $\delta = \delta_0 / (1 + 1/3\,\alpha/\alpha_I)$, where α and α_I are the bond-stretching force constants for the host (BC) and the impurity (AC) crystals. With $\alpha \approx \alpha_I$, this model predicts $\delta = 3/4\delta_0$ for a zinc-blende alloy. Although our full perturbation theory (FPT) and the VFF model of Martin and Zunger (MZ)(4) predict d_{AC} with an average absolute deviation comparable to the experimental uncertainty of 0.01A, the simple spring model is even better.

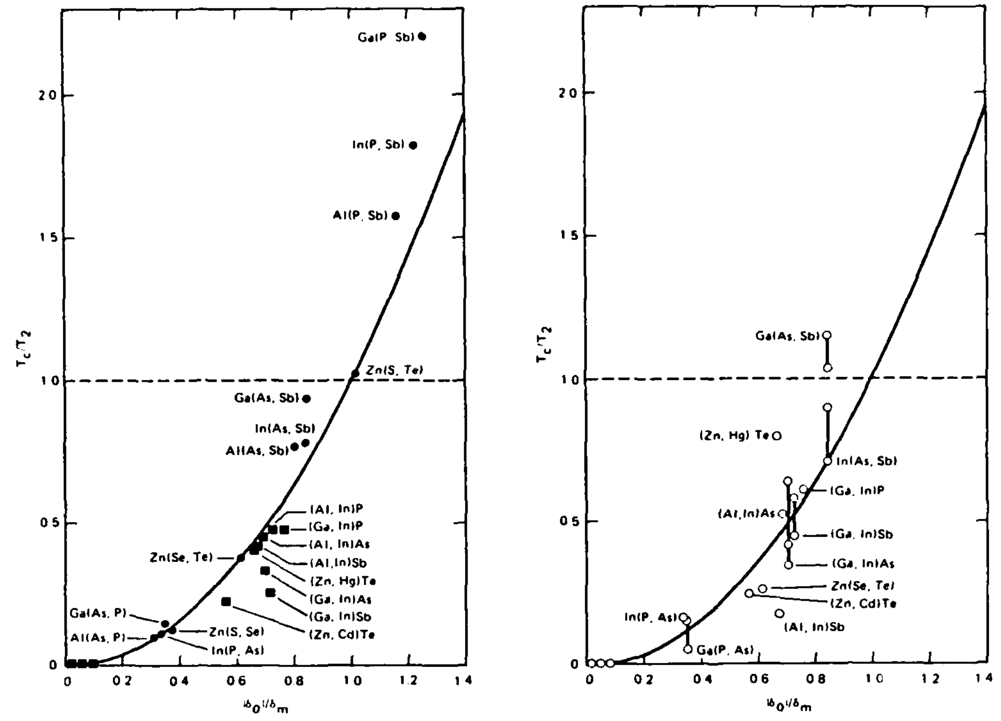
We note that while in our theory, MZ and SSHS, the Ω values are directly calculated without any adjustable parameters, our theory and SSHS agree with the experiment as well or even slightly better than the one-parameter theories.(5,6) Although our theory predicts a negative Ω value for all three (Ga, Al) alloys, the magnitude ($\Omega = -0.17$ kcal/mole) is too small to account for the ordering of $Ga_xAl_{1-x}As$ grown at 600 to 700 °C found recently.(7) The calculated Ω values also provide guidance in separating miscible from immiscible alloys. In a random alloy, the criterion for alloy mixing for all x is $T > T_c$, where the critical temperature T is given by $T_c = \Omega/2R_g$ with R_g being the universal gas constant. Figure 1 is a plot of T_c/T_2 against $|\delta_0| / |\delta_m|$, where T_2 is the lower of the two constituent's melting temperatures, and $\delta_m = 1.63\chi_m$ with χ_m being the ratio of rms bond length amplitude fluctuation to the bond length at T_2. The simple spring model gives $T_c/T_2 = (\delta_0/\delta_m)^2$ as indicated by the solid curves.

There is an empirical rule stating that a miscibility gap will occur if $|\delta_0|$ between two alloy components exceeds 7.5%. However, if T_c/T_2 is plotted against $|\delta_0|$, the simple spring model would not exhibit a smooth simple quadratic curve, and our theoretical points would be much more scattered. This suggests that $|\delta_0| / |\delta_m| > 1$ is a better criterion than $|\delta_0| > 0.075$. Figure 2 also clearly shows the chemical effects, namely negative and positive chemical energies ΔE_{ch} for cation and anion substitutions respectively. The full theory and the experiments correlate within the experimental uncertainties. The simple SSIIS model clearly is an excellent universal representation. However, T_c/T_2 varies faster than quadratically for larger $|\delta_0/\delta_m|$ values, as born out from both the experimental data and the full theory.

CONCENTRATED ALLOYS

Turn now to the concentrated alloy case. First, an improved statistical model is required. We have extended regular solution theory based on pair energies to one for five-atom clusters. For an $A_xB_{1-x}C$ alloy, the building blocks are clusters of A(m)B(4-m)C, where m ranges from 0 to 4. For a given alloy concentration x and for a given set of energies ϵ_m associated with these clusters, we have derived expressions for the cluster population distribution $x_m \equiv \bar{n}_m / N$, where N is the total number of unit cells and $\bar{n}_m$ is the averaged number of cells with A(m)B(4-m)C clusters. The partition function Z is obtained using a steepest descents argument which then yields the mixing Helmhotz free energy ΔF. The result reduces to Guggenheim's tetrahedron case(7) if pair potentials (for the second-neighbors) are assumed. Another major difference is that we only need to solve a single quartic equation, while Guggenheim needed to solve four simultaneous quartic equations.

The key to the problem, however, lies in the calculation of the energies ϵ_m. If one assumes that the size of the tetrahedra for all m-clusters at a given alloy concentration takes on the corresponding VCA values but allows the central C atom to relax, then the energies as functions of x behave like those shown in Figure 2(a). There are at least two major flaws in this result. First, the energies are too large and would correspond to Ω values many times the experimental values. Second, at x = 0.75, 0.5, and 0.2, these energies imply compound formation for $A_3B_1C_4$, $A_2B_2C_4$, and $A_1B_3C_4$, respectively, which is opposite to the known tendency for spinodal decomposition of $Ga_xIn_{1-x}As$ at low T. However, if the local cell volume of each cluster is allowed to be in mechanical equilibrium with a continuous medium with an effective shear coefficient $\overline{C} = xC_{AC}^{(0)} + (1-x)\, C_{AB}^{(0)}$, where the C value for the pure material is given by Eq. (1), then, the corresponding energies ϵ_m as a function of x are given in Figure 2(b), which now yields a reasonable value of mixing enthalpy and correctly predicts the tendency toward spinodal decomposition at low temperature. With this set of energies, one can then calculate the cluster distribution x_m, and compare them with the corresponding values for a random alloy, i.e. $x_m^{(0)} = (4_m)x^m(1-x)^{4-m}$. Figure 3 shows the deviation from randomness $\Delta x_m = x_m - x_m^{(0)}$ as a function of x for four arbitrarily chosen growth



(a) THEORY

(b) EXPERIMENT

FIGURE 1 PLOT OF T_c/T_2 AS A FUNCTION $|\delta_o|/\delta_m$ FOR T_c OBTAINED FROM (a) PRESENT THEORY AND (b) THE EXPERIMENTAL Ω. IN (a), THE DOTS ARE FOR ANION ALLOYS AND CROSSES FOR CATION SUBSTITUTION. THE SOLID LINES IN BOTH PARTS CORRESPOND TO THE Ω BASED ON THE SIMPLE THEORY OF REF. 3. THE DASHED LINES AT $T_c/T_2 = 0$ SEPARATE THE MISCIBLE FROM IMMISCIBLE GROUPS

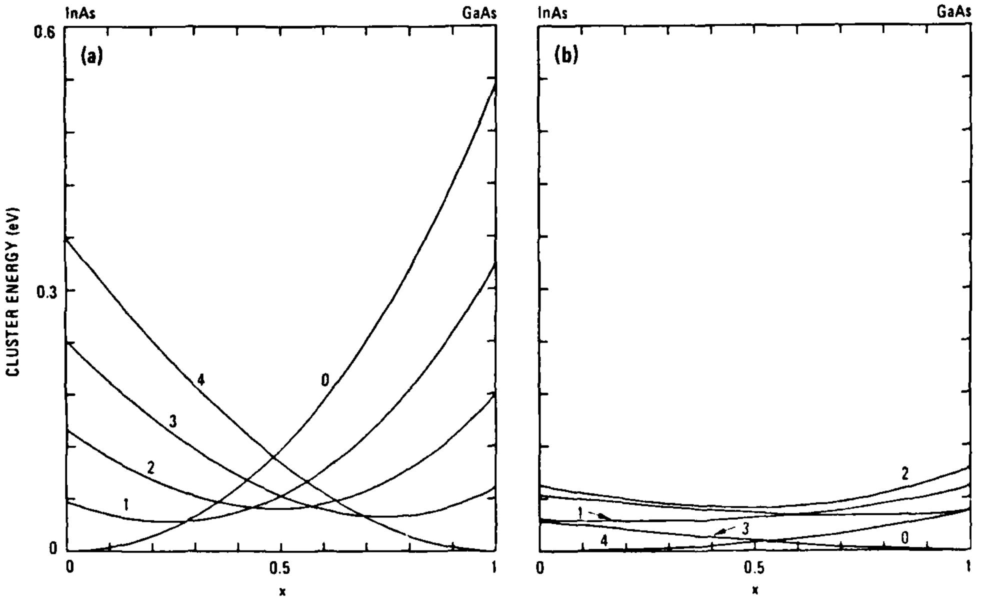


FIGURE 2 THE CLUSTER ENERGIES ϵ_m AS A FUNCTION OF x CALCULATED IN TWO DIFFERENT WAYS, AS DISCUSSED IN THE TEXT

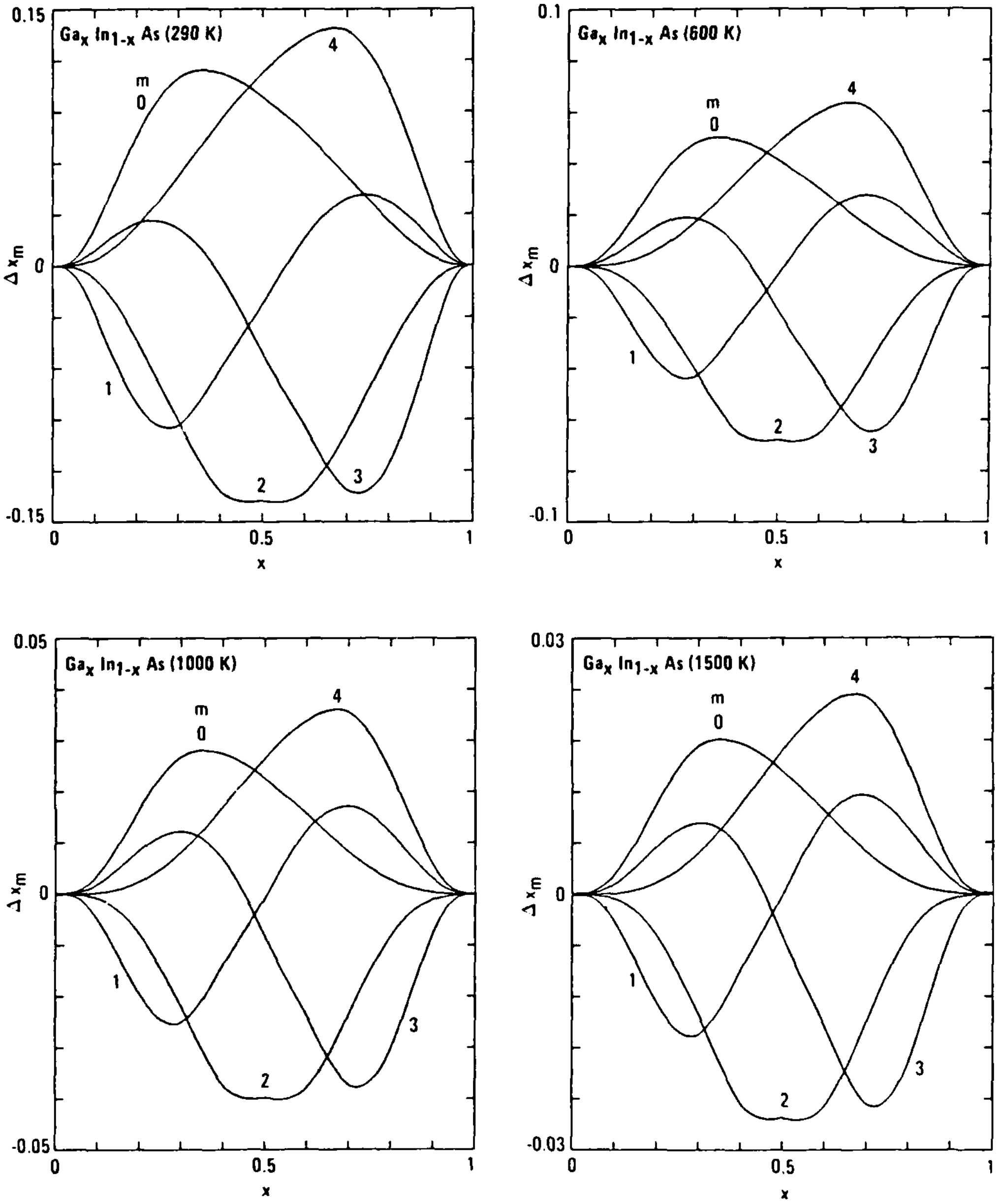


FIGURE 3 DEVIATIONS FROM RANDOM DISTRIBUTIONS FOR THE FIVE CLUSTERS AT SEVERAL GROWTH TEMPERATURES

temperatures. It shows that Δx_m can be several percent at ordinary growth temperatures and there are deviations from the symmetric behavior about $x = 1/2$ predicted by the pair energy model.

It is also interesting to examine the sets of bond lengths $d_{AC}^{(m)}$ and $d_{BC}^{(m)}$ corresponding to the ϵ_m values of Figure 2(b). The spread in the d values among different clusters is found to lie within the width of EXAFS lines.

The existence of microclusters, whose populations deviate from those of a random alloy, will impact on many experimental results. These include phase diagrams, EXAFS line positions and widths, magnetic quadruple splittings, modulation spectroscopy widths, Raman frequencies and intensities for different modes, etc. We have demonstrated that the theory agrees with EXAFS and measured mixing enthalpy parameters. It should be tested against more experiments to further our understanding of the structural properties of semiconductor alloys.

Acknowledgment

The work is supported in part by AFOSR-84-0284 and DARPA Contract MDA 903-83-C-0108. A.-B. Chen would like to thank Professor W.E. Spicer for his hospitality at Stanford University.

REFERENCES

1. J.C. Mikkelsen and J.B. Boyce, Phys. Rev. Lett. *49*, 1412 (1982); Phys. Rev. *28*, 7130 (1983); and private communications.
2. R.M. Martin, Phys. Rev. *B1*, 4005 (1970).
3. C.K. Shih, W.E. Spicer, W.A. Harrison and A.Sher, Phys Rev. B*31*, 1139 (1985).
4. J.L. Martins and A. Junger, Phys. Rev. B *30*, 6217 (1984).
5. G.B. Stringfellow, J. Cryst. Growth *27*, 21 (1974); *58*, 194 (1982).
6. P.A. Fedders and M.W. Muller, J. Phys. Chem. Solids *45*, 685 (1984).
7. T.S. Kuan, T.F. Kuech, W.I. Wang, and E.L. Wilkie, Phys. Rev. Lett. *54*, 201 (1985).
8. E.A. Guggenheim, *Mixtures*, (Oxford at the Clarendon Press, 1952).

SEMICONDUCTOR PSEUDO-BINARY ALLOYS:
BOND-LENGTH RELAXATION AND MIXING ENTHALPIES

A.-B. Chen,* Physics Department
Auburn University, Alabama 36849
A. Sher, SRI International
Menlo Park, California 94025

Harrison's bonding theory, the valence force field (VFF), and an elastic continuum are combined to study the substitution energies Δ_s and local bond lengths d_1 of isoelectronic impurities in semiconductors. Explicit expressions for Δ_s and d_1 are derived, which enable us to absorb measured elastic constants into the calculation and to study the chemical effects arising from differences in the covalent radii and polarities. Several models based on VFF alone are also derived for comparison. The full theory and at least five VFF models are found to produce impurity bond lengths in excellent agreement with experiment. The substitution energies are shown to provide good estimates of the mixing enthalpies Ω of pseudo-binary alloys and to predict miscibility gaps properly. The chemical shifts in Ω are found to be negative for most cation alloys but positive for anion substitutions.

I INTRODUCTION

The discovery of a bimodal distribution of the nearest-neighbor bond lengths[1] in $Ga_xIn_{1-x}As$ has sparked considerable interest in the bonding nature of semiconductor alloys.[2-6] This finding has changed the conventional picture of the alloy crystal bond configuration, which has far-reaching implications about the electronic structure, structural stability, and thermodynamics of these materials. Because of the complexity of both the structural and the potential disorder in these alloys, ab-initio band-structure techniques have not yet evolved to a stage suitable for direct calculations. Therefore, we have extended Harrison's bonding theory[7] to study the alloy structural properties.[5,6] In this paper, we apply an intermediate version of the theory to the dilute-limit case of an isoelectronic impurity.

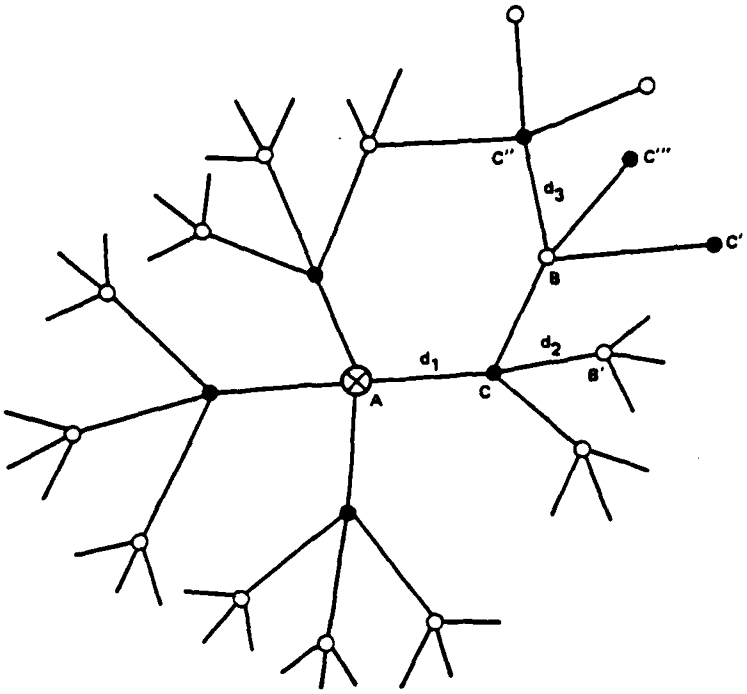
A particularly useful application of the theory is its perturbation-expansion form, in which measured elastic constants are incorporated to obtain accurate results. This form is also useful for comparison with other previously published models[3,8,9] that are based on the valence force field (VFF)[10] model alone. Thus, all the factors influencing bond-length relaxation, e.g. strains, boundary conditions, and chemical effects, can be studied. The ability to incorporate the chemical effects is one major difference between this theory and other VFF models.

The rest of the paper contains the following sections: Section II describes a theory for calculating impurity substitution energies. Section III casts the theory into perturbation form and combines it with a valence force field and an elastic continuum. Several VFF models are derived in Section IV. The modifications of numerical results due to chemical effects on local bond lengths and alloy mixing enthalpies are summarized and discussed in Section V. Conclusions are drawn in the last section, Section VI.

II. IMPURITY SUBSTITUTION ENERGY

Consider the problem of substituting an isoelectronic atom A for a B atom in a zinc-blende compound BC (e.g., In substitutes for Ga in GaAs, as shown in Figure 1). In general, the bond lengths d_1, d_2, d_3 ... for the first-, second-, and third-shell bonds surrounding the impurity are different from the equilibrium values of either the pure BC compound, denoted as d, or the "impurity" compound AC, denoted as $d_I = d\,(1-\delta_o)$. If A starts being a free atom, and B also ends being a free atom, then the energy difference between the final and initial states is defined as the substitution energy and is given by

$$\Delta_s = (E_{def}+\varepsilon_B)-(E_{pure}+\varepsilon_A), \qquad (1)$$



	POSITIONS	DISPLACEMENT
C:	$\frac{d}{\sqrt{3}}(1, 1, 1)$	$\frac{d}{\sqrt{3}}(-\delta, -\delta, -\delta)$
B:	$\frac{d}{\sqrt{3}}(0, 2, 2)$	$\frac{d}{\sqrt{3}}(0, \gamma, \gamma)$
B':	$\frac{d}{\sqrt{3}}(2, 0, 2)$	$\frac{d}{\sqrt{3}}(\gamma, 0, \gamma)$
C':	$\frac{d}{\sqrt{3}}(1, 3, 3)$	$\frac{d}{\sqrt{3}}(\gamma', 3\gamma', 3\gamma')$
C'':	$\frac{d}{\sqrt{3}}(-1, 3, 1)$	$\frac{d}{\sqrt{3}}(-\gamma'', 3\gamma'', \gamma'')$
C''':	$\frac{d}{\sqrt{3}}(-1, 1, 3)$	$\frac{d}{\sqrt{3}}(-\gamma'', \gamma'', 3\gamma'')$

Fig. 1. A sketch of the flatted picture of a single impurity A in a BC compound. The positions and displacements for those atoms labeled are used in Appendix B.

where ϵ_A and ϵ_B are free-atom energies for A and B respectively, E_{def} is the total energy of the semiconductor with a defect as sketched in Figure 1, and E_{pure} is that of the pure host BC crystal. Equation (1) can be written as

$$\Delta_s = (E_{def}-E_{dis})-(\epsilon_A-\epsilon_B)+(E_{dis}-E_{pure})$$

$$= \Delta_R-(\epsilon_A-\epsilon_B)+\Delta_{dis}, \qquad (2)$$

where we have added and substracted a term E_{dis}, which is the total energy of a distorted BC compound with all the atoms held at the positions specified in Figure 1, except the central atom is a B atom. Clearly, $\Delta_R = E_{def}-E_{dis}$ is a replacement energy, and the distortion energy $\Delta_{dis} = E_{dis}-E_{pure}$ is the energy required to deform a pure BC crystal from its equilibrium lattice to that specified in Figure 1. Δ_R contains all the chemical terms that arise from different bond lengths and polarities between AC and BC bonds.

Δ_R can be treated most easily by Harrison's bonding theory.[7] In this theory, the energy per bond relative to the vacuum state is

$$E_b = 2\epsilon_b+V_o+6\epsilon_m^{+}+6\epsilon_m^{-}, \qquad (3)$$

where ϵ_b is the energy of the bonding state constructed from the two hybrid orbitals facing each other along the bond direction

$$E_b = 1/2\ (\varepsilon_h^A+\varepsilon_h^C)-(V_2^2+V_3^2)^{1/2}, \qquad (4)$$

with $\varepsilon_h{}^A$ and $\varepsilon_h{}^C$ being the energies of the anion and cation hybrids orbitals respectively. The antibonding energy ε_a has the same form as in Eq.(4) except with a plus sign. V_2 is called the covalent energy, which is the total electronic Hamiltonian matrix element between the two hybrids in question, and the polar energy V_3 is the difference $V_3 = 1/2\ (\varepsilon_h{}^C-\varepsilon_h{}^A)$. The $\varepsilon_m{}^+$ and $\varepsilon_m{}^-$, the metallization energies, are the shifts of the bonding level caused by interactions with the neighboring anti-bonding states, where + and - indicate whether the common adjacent atom is a cation or anion. For example, this term for an AC bond labeled by d_1 in Fig. 1 due to an antibonding state labeled by d_2 is given by

$$\varepsilon_{m,C}^{B,A}(2,1) = |U_b^A(1)|^2|U_a^B(2)|^2\ |V_1^C|^2/[\varepsilon_b^A(1)-\varepsilon_a^B(2)], \qquad (5)$$

where A and B denote AC and BC bonds respectively, $V_1{}^C = 1/4\ (\varepsilon_s{}^C-\varepsilon_p{}^C)$ with $\varepsilon_s{}^C$ and $\varepsilon_p{}^C$ being the s- and p- term values of the common adjacent atom C. $U_b{}^A(1)$ is the probability amplitude of finding an electron in the hybrid orbital of the C atom in the bonding state of an AC bond with a bond length d_1, whereas $U_a{}^B(2)$ is the corresponding probability amplitude for the antibonding state of a BC bond of bond length d_2. Finally V_o is a repulsive pair potential required to prevent the crystal from collapsing and to guarantee a correct equilibrium bond length.

The local perturbation, Eq.(5), is applicable, because the square root of the numerator is much smaller than the separation between the antibonding and bonding levels, and the valence band is completely filled, so the interaction between the bonding states only spreads the ε_b levels into bands without affecting the center of gravity of the occupied states.

The replacement energy Δ_R of Eq.(2) can now be written explicitly:

$$\begin{aligned}\Delta_R = 4[&2\varepsilon_b^A(1)+V_o^A(1)+6\varepsilon_{m,A}^{A,A}(1,1)+6\varepsilon_{m,C}^{B,A}(2,1)\\ &-2\varepsilon_b^B(1)-V_o^B(1)-6\varepsilon_{m,B}^{B,B}(1,1)-6\varepsilon_{m,C}^{B,B}(2,1)\\ &+6\varepsilon_{mC}^{A,B}(1,2)-6\varepsilon_{m,C}^{B,B}(1,2)].\end{aligned} \tag{6}$$

The distortion energy Δ_{dis} of Eq.(2) now involves only BC bonds of different bond lengths. It can be treated with exactly the same procedure for any given set of bond-length distributions. Thus, a straightforward energy minimization procedure can be carried out. The accuracy of this procedure, however, depends in turn on the accuracy of scaling rules for V_2 and V_o and the input parameters. At present, Harrison's model[7] with $V_2 \propto 1/d^2$ and $V_o \propto 1/d^4$ and his universal parameters are only semi-quantitative. We are improving the quantitative nature of the theory so that the full theory will yield accurate predictions of the structural and thermal properties of semiconductor defects and alloys.

III. PERTURBATION EXPANSION, VALENCE FORCE FIELD, AND ELASTIC MEDIUM

As pointed out earlier, a perturbation expansion of the theory is instructive. This is feasible because the differences $\Delta V_2 = V_2(AC)-V_2(BC)$ and $\Delta V_3 = V_3(AC)-V_3(BC)$ are small compared to each individual value for many of the isoelectronic impurities in III-IV and II-VI compounds. To this end, Eq.(6) is rewritten as

$$\begin{aligned}(1/4)\Delta_R = {} & E_b^A(1)-E_b^B(1)+6[\varepsilon_{m,C}^{B,A}(2,1)-\varepsilon_{m,C}^{A,A}(1,1)] \\ & -6[\varepsilon_{m,C}^{A,B}(2,1)-\varepsilon_{m,C}^{B,B}(1,1)]+6[\varepsilon_{m,C}^{A,B}(1,2)-\varepsilon_{m,C}^{B,B}(2,2)] \\ & -6[\varepsilon_{m,C}^{B,B}(1,2)-\varepsilon_{m,C}^{B,B}(2,2)], \qquad (7)\end{aligned}$$

where $E_b{}^A(1)$ and $E_b{}^B(1)$ are energies per bond in Eq.(3) for AC and BC compounds respectively, with the relaxed bond length $d_1 = d\,(1-\delta)$. The difference between these energies E_b and the corresponding values at their respective equilibrium bond lengths d_I and d are just the strain energies per bond in uniform deformation:

$$\begin{aligned}E_b^A(1) &= E_b^A(d_I)+2\sqrt{3}\ B_I d_I(d_1-d_I)^2, \\ E_b^B(1) &= E_b^B(d)+2\sqrt{3}\ Bd(d_1-d)^2, \qquad (8)\end{aligned}$$

where B_I and B are the bulk moduli for the "impurity" AC and host BC crystals. The rest of the terms in Eq.(7) are all due to changes in ε_m caused by the differences ΔV_3 and ΔV_2. We shall use Harrison's scaling rules to deduce them.[7] Expanding Eq.(7) to second order in ΔV_3 and $\Delta d = d_2 - d_1$, we write

$$6[\varepsilon_{m,C}^{B,A}(2,1)-\varepsilon_{m,C}^{A,A}(1,1)] = f_I\Delta d-g_I\ \Delta V_3+h_I(\Delta d)^2$$
$$-W_I\Delta d\Delta V_3+U_I(\Delta V_3)^2, \qquad (9)$$

where f_I, g_I, and so on, are appropriate derivatives evaluated for the "impurity" crystal AC. When similar expansions are made for the rest of the terms in Eq.(7), it becomes [with $d_1 = d(1-\delta)$, $d_I = d(1-\delta_o)$]

$$(1/4)\Delta_R = \Delta E_b+(f_I-f)\Delta d-(g_I-g)\Delta V_3+(h_I+h)(\Delta d)^2$$
$$-(W_I+W)\ \Delta d\Delta V_3+(U_I+U)(\Delta V_3)^2$$
$$+2\sqrt{3}B_I d^3(\delta-\delta_o)^2-2\sqrt{3}\ Bd^3\delta^2, \qquad (10)$$

where

$$\Delta E_b = E_b^A(d_I)-E_b^B(d)-(1/4)(\varepsilon_A-\varepsilon_B) \qquad (11)$$

is just the difference in the binding energy per bond between the "BC" and "AC" crystals. In Eq.(10), the coefficients f, g, without a subscript are those for the host BC system. It is convenient to define an excess energy $\Delta E = \Delta_s/4-\Delta E_b$, which is the extra energy per bond required for the impurity substitution over and above the binding energy difference between the BC and AC crystals. The binding energy difference accounts for much of the substitution energy; however, the correction measured by the excess energy can be significant. The excess energy results from strain energies and chemically driven charge redistributions around the defect. Using Eqs.(2) and (10) and defining $F = f_I-f$ and $G = g_I-g$, we can write ΔE up to second order in ΔV_3 and Δd as

$$\Delta E = 2\sqrt{3}\ B_I d^3(\delta-\delta_o)^2-2\sqrt{3}\ Bd^3\delta^2+F\Delta d-G\Delta V_3$$
$$+H(\Delta d)^2+W\Delta d\Delta V_3+U(\Delta V_3)^2+(1/4)\Delta_{dis}, \tag{12}$$

where $H = h_I + h$, $W = w_I + w$, and $U = u_I + u$.

To treat the distortion energy Δ_{dis}, we divide the crystal into two regions. Inside a sphere of some radius R measured from the impurity, the strain energy is taken to be the valence force field[10] value:

$$\Delta_{dis}^{(in)} = \frac{3}{8d^2}\sum_i \alpha_i[\Delta(\vec{d}_i\cdot\vec{d}_i)]^2+\frac{3}{8d^2}\sum_{i>j}{}'\sum \beta_{ij}[\Delta(\vec{d}_i\cdot\vec{d}_j)]^2, \tag{13}$$

where i sums over all the bonds inside R, and the pairs in the β-terms include those that have adjacent atoms inside R and on the boundary. The parameters α and β are force constants to be considered later. $\Delta(\vec{d}\cdot\vec{d}_j) = \vec{d}_i\vec{d}_j-\vec{d}_i^{\,(0)}\cdot\vec{d}_j^{\,(0)}$ measures the change of the dot product between bond vectors due to distortions. Outside R, we assume an elastic continuum with radial displacements inversely proportional to the square of the distance from the center. The elastic energy in this medium can be shown to be (see Appendix A)

$$\Delta_{dis}^{(out)} = RCu^2 \qquad (14)$$

where the effective shear coefficient is $C = 4\pi[0.4(C_{11}-C_{12})+1.2\ C_{44}]$, and u is the displacement at R. In view of the fact that the bonds d_1 and d_2 are coupled through the chemical terms in Eq.(10), the smallest logical radius R is the second-shell atomic distance, namely $R = 2\sqrt{2}d/\sqrt{3}$. Atoms on this boundary have displacements of the forms $\vec{u} = d(\gamma,\gamma,0)/\sqrt{3}$, ... Thus, $u = \sqrt{2}\ \gamma\ d/\sqrt{3}$ and the elastic energy in the continuum is

$$\Delta_{dis}^{(out)} = \frac{4}{3}\sqrt{\frac{2}{3}}\,C\gamma^2 d^3. \qquad (15)$$

The distortion energy represented by Eq.(14) contains six different contributions (see Appendix B): the bond-stretching energy of the four first-shell bonds, $6\alpha\delta^2d^2$, the β-terms from the first-shell bonds, $\beta\delta^2d^2$, the α-terms from the second-shell bonds, $2\alpha(\delta+2\gamma)^2d^2$, the β-terms between the first-and second-shell bonds, $2\beta(\delta+\gamma)^2d^2$, the β-terms among the second-shell bonds $2\beta\delta^2d^2$, and finally the β-terms between the second-shell bonds and those in the continuum, $(1/2)\beta d^2[(3\delta+\lambda_1\gamma)+(\delta+\lambda_2\gamma)^2]$, where $\lambda_1 = 40\sqrt{2}/(19\sqrt{19})$ and $\lambda_2 = 2-40\sqrt{2}/(11\sqrt{11})$.

To assemble all the contributions to Eq.(12), we need to consider the assignments of the elastic constants and the force constants α and β in VFF. While the experimental values[11] of C_{11}, C_{12}, and C_{44} can be used for the elastic constants, α and β have to be deduced. If Martin's original procedure[10] (also followed by Martins and Zunger[3]) is used, then Eq.(13) alone will not produce the correct (experimental) bulk moduli. There are small corrections due to Madelung terms, which are hard to treat in the case of non-uniform distortions. A simpler procedure is adopted here. We use the experimental bulk moduli for B_I and B in Eq.(12) and experimental elastic constants to calculate C of Eq.(14), and then force α and β in the VFF to produce the correct bulk moduli B and shear coefficients $C_{11}-C_{12}$. Such an approach is also consistent with Harrison's bonding theory[7] and other approaches in which the Coulomb forces are automatically incorporated in the band and bond energies, and do not need to be redundantly treated. With our procedure, the bulk modulus is simply related to the force constants by $B = (3\alpha + \beta)/(4\sqrt{3}d)$. Table 1 lists our α and β values. We want to point out in advance, however, that the numerical results deduced from our and Martin's sets of α and β do not introduce differences more than the present experimental uncertainties in the local bond length (~ 0.01Å) and the mixing enthalpies (> 0.5 kcal/mole).

Table 1. The bond lengths d (in Å, valence force constants α and β (in N/m), shear coefficients C of the continuum (in 10^{11} erg/cm^3), melting temperatures T_m(K), and Liedermann ratios X_m for the compounds used in this paper.

Compound	d	α	β	C	T_m*	X_m
AlP	2.367	44.323	8.068	122.396	1773	0.070
AlAs	2.451	40.849	8.717	112.695	1873	0.073
AlSb	2.656	34.073	6.900	85.351	1323	0.062
GaP	2.360	44.764	10.737	145.921	1510	0.064
GaAs	2.448	39.235	9.159	121.844	1738	0.071
GaSb	2.640	31.876	7.347	89.372	985	0.055
InP	2.541	40.363	6.543	91.785	1343	0.059
InAs	2.622	33.203	5.752	78.816	1215	0.061
InSb	2.805	28.557	4.891	60.721	798	0.049
ZnS	2.342	40.429	5.273	89.272	2123	0.081
ZnSe	2.454	32.200	4.562	82.687	1788	0.080
ZnTe	2.637	29.445	4.659	62.430	1511	0.071
CdTe	2.806	26.569	2.722	38.453	1371	0.067
HgTe	2.798	26.396	2.746	40.363	943	0.056

---

*Ref. 30

Using the above procedure and adding all contributions, the excess energy per bond from Eqs.(10) and (12) is the full perturbation theory (FPT) result

$$\Delta E = [3\alpha_I(\delta-\delta_o)^2/2+\beta_I(\delta-\delta_o)^2/2+\alpha(\delta+2\gamma)^2/2$$
$$+\beta(\delta+\gamma)^2/2+\beta\delta^2/4+\beta(3\delta+\lambda_1\gamma)^2/8+\beta(\delta+\lambda_2\gamma)^2/4]d^2$$
$$+\sqrt{2}\ C\gamma^2d^3/(3\sqrt{3})+\Delta E_{ch}, \qquad (16)$$

where the chemical contribution is written as

$$\Delta E_{ch} = F_{ch}\Delta d+H(\Delta d)^2+\Delta E_p$$
$$= \frac{2}{3}\ F_{ch}(2\delta+\gamma)d+\frac{4}{9}\ H(2\delta+\gamma)^2d^2+(U\Delta V_3^2\ -G\Delta V_3), \qquad (17)$$

where $\Delta F_{ch} = F-W\Delta V_3$ and $\Delta E_p = U\Delta V_3^2-G\Delta V_3$. F_{ch} is a chemical force, which, when it is positive, tends to push the C atom away from the impurity atom A. This force arises from the difference in the bond tensions induced between the AC and BC bonds adjacent to C because the neighboring anti-binding states are different from those of their respective host states. ΔE_p is due to the difference in the polarities ΔV_3 alone and is independent of the displacement. Finally H can be regarded as a chemically induced force constant which, when it is positive, tends to restrain the lattice from distortion and increases the elastic energy.

The equilibrium requirements, $\partial(\Delta E)/\partial\delta = 0$ and $\partial(\Delta E)/\partial\gamma = 0$, then lead to the solution $\gamma = Q\delta$, and δ is given by

$$\delta = (\delta + \delta_o')/\{1+[\alpha(1-2Q)+B(17/4-\lambda Q)+16H(1-2Q)/9]/(3\alpha_I+\beta_I)\}. \tag{18}$$

where the constant λ is $1+3\lambda_1/4+\lambda_2/2$, and

$$\delta_o' = -4F_{ch}(1-Q/2)/[3d(3\alpha_I+\beta_I)], \tag{19}$$

with $Q = 2J/K$, $J = \alpha+\lambda\beta/2+8H/9$, and K

$$= 4\alpha+2\sqrt{2}Cd/(3\sqrt{3})+(1+\lambda_1^2/4+\lambda_2^2/2)\beta+8H/9.$$

IV. VALENCE FORCE FIELD MODELS

In this section we consider several models based on the valence force field (VFF). These models have been used frequently to explain the impurity bond relaxation [3,4,8,9]. We shall first derive the explicit expressions for these models and then connect them with the existing results.

Model A: Third-Shell Atoms and Beyond Are Fixed at Their Pure Crystal Positions

Let the bond lengths surrounding the impurity again be $d_1 = d(1-\delta)$ and let the second-shell atoms have radial displacements of the forms $(d/\sqrt{3})(\gamma,\gamma,o)$ etc. Beyond and including the third-shell, all the other atoms are held at their pure-crystal positions. There are nine different contribubutions to the strain energy in VFF (see Appendix B): the α-terms from the four bonds surrounding the impurity: $6\alpha_I(\delta-\delta_o)^2d^2$, the β-terms among the six pairs of these bonds, $\beta_I(\delta-\delta_o)^2d^2$, the β-terms between the first-shell and second-shell bonds, $2\beta(\delta+\gamma)^2d^2$, the α-terms from the second-shell bonds, $2\alpha(\delta+2\gamma)^2d^2$, the β-terms among the second-shell bonds, $2\beta\delta^2d^2$, the β-terms between the second-shell and third-shell bonds: $(9/2)\beta\delta^2d^2 + \beta(\delta+2\gamma)^2d^2$, the α-terms from the third-shell bonds, $8\alpha\gamma^2d^2$, the β-terms among the third-shell bonds, $4\beta\gamma^2d^2$, and the β-terms between the third and fourth-shell bonds $6\beta\gamma^2d^2$. Thus, the excess energy (in this case 1/4 times the strain energy) becomes

$$\Delta E = [\tfrac{3}{2}\alpha_I(\delta-\delta_o)^2+\tfrac{1}{4}\beta_I(\delta-\delta_o)^2+\tfrac{1}{2}\beta(\delta+\gamma)^2 + \tfrac{1}{2}\alpha(\delta+2\gamma)^2+\tfrac{13}{8}\beta\delta^2+\tfrac{1}{4}\beta(\delta+2\gamma)^2 + 2\alpha\gamma^2+\tfrac{5}{2}\beta\gamma^2]d^2. \tag{20}$$

The minimization of ΔE w.r.t δ and γ leads Eq.(20) to $\gamma = -\delta/4$, and

$$\delta = \delta_o/[1+(\alpha+17\beta/2)/(6\alpha_I+\beta_I)]. \tag{21}$$

We note that there is some ambiguity in the third contribution listed above for the β-terms between the impurity and host bonds. The value of β could be chosen as one of these combinations β, β_I, $1/2\ (\beta+\beta_I)$, $\sqrt{\beta\beta_I}$ or other proper combinations. Because the values of β and β_I are comparable and β values are much smaller than α (see Table 1), the results for δ and ΔE are not too sensitive to the choice. There is also some ambiguity in the values for $\vec{d}_i^{(0)}\cdot\vec{d}_j^{(0)}$ for the "undistorted" crystal. The $-d^2/3$ used is the simplest choice. A different choice will not affect the results for δ at all, but will make ΔE slightly different. In fact, Model A was firstly used by Martins and Zunger.[3] However, their expression for δ is different from Eq.(21), because they made different choices of the two quantities just mentioned. Nevertheless, Section IV will show that these two expressions yield very similar results. These ambiguities do not occur in the full theory in Section III, where the impurity-host interactions are taken into account naturally by the replacement energy Δ_R [see Eq.(10)].

Model B: Second-Shell Atoms Connect to a Fixed Boundary

This model corresponds to $\gamma = 0$ in Model A. So we have

$$\Delta E = [\frac{3}{2}\alpha_I(\delta-\delta_o)^2+\frac{1}{4}\beta_I(\delta-\delta_o)^2+\frac{1}{2}\alpha\delta^2+\frac{19}{8}\beta\delta^2]d^2, \tag{22}$$

and

$$\delta = \delta_o/[1+(\alpha+19\beta/4)/(3\alpha_I+\frac{1}{2}\beta_I)]. \tag{23}$$

This expression will be used to study the effect of truncation.

Model C: Simple Spring Model

If all the β's in Eqs.(22) and (23) are set equal to zero, we have the simple spring model with

$$\Delta E = [\frac{3}{2}\alpha_I(\delta-\delta_o)^2+\frac{1}{2}\alpha\delta^2]d^2 \tag{24}$$

and

$$\delta = \delta_o/(1+\frac{1}{3}\alpha/\alpha_I). \tag{25}$$

The spring model recently[8] discussed by Shih et al. corresponds to Eq.(25) with $\alpha = \alpha_I$, so $\delta/\delta_o = 3/4$.

Model D. VFF With the Continuum Connected to the Second-Shell Atoms

Model D1

In this case, ΔE only contains the first five contributions listed for Case A plus the elastic energy in the continuum. However, the β-terms between the second- and third-shell bonds are modified because atoms outside R in the continuum now have radial displacements proportional to the inverse of the square of the radius. The result is

$$\Delta E = [\tfrac{3}{2}\alpha_I(\delta-\delta_o)^2+\tfrac{1}{4}\beta_I(\delta-\delta_o)^2+\tfrac{1}{2}\alpha(\delta+2\gamma)^2+\tfrac{1}{2}\beta(\delta+\gamma)^2$$

$$+\tfrac{1}{2}\beta\delta^2+\tfrac{1}{8}\beta(3\delta+\lambda_1\gamma)^2+\tfrac{1}{4}\beta(\delta+\lambda_2\gamma)^2+\frac{\sqrt{2}}{3\sqrt{3}}Cd\gamma^2]d^2, \qquad (26)$$

where λ_1 and λ_2 are the same as the constants that appear in Eq.(16). The corresponding equilibrium condition can be shown to be

$$\delta = \delta_o/\{1+[\alpha(1-2Q)+19\beta/4-\beta(1+3\lambda_1/4+\lambda_2/4)Q]/(3\alpha_I+\beta_I/2), \qquad (27)$$

where

$$Q = (2\alpha+\beta+3\lambda_1\beta/4+\lambda_2\beta/2)/(4\alpha+2\sqrt{2}Cd/3\sqrt{3}+\beta+\frac{1}{4}\lambda_1^2\beta+\frac{1}{2}\lambda_2^2\beta). \tag{28}$$

Model D2

A comparison between Eq.(26) and the full perturbation (FPT) theory Eq.(16) shows two major differences. First, all the chemical terms are absent in Eq.(26). Secondly, the terms $(1/2)\beta_I(\delta-\delta_o)^2d^2+(1/4)\beta\ \delta^2d^2$ in Eq.(16) become $(1/4)\ \beta_I\ (\delta\ -\delta_o)^2d^2+(1/2)\ \beta\ \delta^2d^2$ in Eq.(26). This difference in the strain energy will mask the true effects of chemical forces if δ from Eq.(26) is compared with FPT. A better way to study the chemical effects is to use the following equation:

$$\begin{aligned}\Delta E = [&\tfrac{3}{2}\alpha_I(\delta-\delta)^2+\tfrac{1}{2}\beta_I(\delta-\delta_o)^2+\tfrac{1}{2}\ \alpha(\delta+2\gamma)^2+\tfrac{1}{2}\beta(\delta+\gamma)^2\\ &+\tfrac{1}{4}\beta\delta^2+\tfrac{1}{8}\beta(3\delta+\lambda_1\gamma)^2+\tfrac{1}{4}\beta(\delta+\lambda_2\gamma)^2]d^2\\ &+\frac{\sqrt{2}}{3\sqrt{3}}\ C\gamma^2d^3,\end{aligned} \tag{29}$$

which is Eq.(16) with all the chemical terms neglected. The corresponding δ becomes

$$\delta = \delta_o/\{1+[\alpha(1-2Q)+\frac{17}{4}\beta-\beta(1+\frac{3}{4}\lambda_1+\frac{1}{4}\lambda_2)Q]/(3\alpha_I+\beta_I)\}, \tag{30}$$

with Q still given by Eq.(28).

Model E: Continuum Connected to the First-Shell Atoms

In this case $\gamma = -3\sqrt{3}\,\delta/(8\sqrt{2})$ and ΔE only includes the first three contributions listed in Model A plus the strain energy of the continuum:

$$\Delta E = [\tfrac{3}{2}\alpha_I(\delta-\delta_o)^2+\tfrac{1}{4}\beta_I(\delta-\delta_o)^2+\tfrac{1}{2}\beta(1-\tfrac{3\sqrt{3}}{8\sqrt{2}})^2\gamma^2 + \tfrac{1}{4}Cd\delta^2]d^2. \tag{31}$$

The relaxation parameter is given by

$$\delta = \delta_o/\{1+[\tfrac{1}{2}\ Cd+(1-\tfrac{3\sqrt{3}}{8\sqrt{2}})^2\beta]/(3\alpha_I+\tfrac{1}{2}\beta_I)\}. \tag{32}$$

We note that the continuum model used to estimate the bond-length relaxation by Baldereschi and Hoffield[9] corresponds to Eq.(32) without the β terms, which yields $\delta/\delta_o = 0.4$ to 0.5, rather than the proper values around 0.7 to 0.8.

V. ALLOY MIXING ENTHALPY

The impurity substitution excess energies ΔE provide a first estimate of the mixing enthalpies of pseudo-binary alloys. Most current thermodynamics theories of semiconductor alloys are based on an extension of the binary solution model.[12] In this model, the mixing Helmholtz energy of an $A_xB_{1-x}C$ alloy is defined as

$$\Delta F_m = F_{alloy}-(xF_{AC}+yF_{BC}), \tag{33}$$

where $y = 1-x$, and F_{AC} and F_{BC} are the respective free energies of the pure AC and BC compounds at the same temperature. Because the C atoms occupy a sublattice, the nearest neighbors of A and B atoms in the alloy are the C atoms. Thus the pair potentials that enter the binary solution theory are now the second-neighbor interactions. Let N_{AA}, N_{AB}, and N_{BB} be respectively the numbers of the second-neighbor AA, AB, and BB pairs, with corresponding pair interaction energies, ε_{AA}, ε_{AB}, and ε_{BB}. For tetrahedral semiconductors, there are a total of 6N second-neighbor pairs for a crystal containing N unit cells. Denote the ratios N_{AA}, N_{AB}, and N_{BB} to 6N as r_{AA}, $r_{AB} = r$, and r_{BB} respectively. Then those ratios are related to the alloy composition by $r_{AA} = x-r/2$ and $r_{BB} = y-r/2$. The mixing free energy has two terms

$$\Delta F_m = \Delta E_m - T\Delta S,$$

where the mixing energy is given by

$$\begin{aligned}\Delta E_m &= E_{alloy}-(xE_{AC}+yE_{BC})\\ &= 6N(\varepsilon_{AA}r_{AA}+\varepsilon_{AB}r_{AB}+\varepsilon_{BB}r_{BB})-6N(x\varepsilon_{AA}+y\varepsilon_{BB})\\ &= 6Nr\Delta\varepsilon, \qquad (35)\end{aligned}$$

where

$$\Delta\varepsilon = \varepsilon_{AB}-\frac{1}{2}(\varepsilon_{AA}+\varepsilon_{BB}). \qquad (36)$$

The mixing entropy ΔS can also be written from a simple generalization of the random distribution.[12] For modest pressure, ΔE is the same as the mixing enthalpy ΔH_m.

Now the pair interaction energies can be approximately related to the impurity substitution energies by

$$\Delta_s(\text{A in BC}) = 12(\varepsilon_{AB}-\varepsilon_{BB}), \qquad (37)$$

and

$$\Delta_s(\text{A in BC}) \approx 12\ (\varepsilon_{AB}-\varepsilon_{AA}). \qquad (38)$$

Thus $\Delta\varepsilon$ of Eq.(36) becomes

$$\Delta\varepsilon = \frac{1}{24}[\Delta_s(\text{A in BC})+\Delta_s(\text{B in AC})]$$

$$= \frac{1}{6}\,[\Delta E(\text{A in BC})+\Delta E(\text{B in AC})]. \tag{39}$$

Usually, the experimental ΔH_m is written as

$$\Delta H_m = x(1-x)\Omega, \tag{40}$$

which is equivalent to assuming a random distribution, i.e. $r = 2x(1-x)$. Using this expression for r and comparing Eq.(40) and (35), we see that the mixing enthalpy parameter Ω is given by

$$\Omega = 2[\Delta E(\text{A in BC}) + \Delta E(\text{B in AC})]. \tag{41}$$

This connection provides a further check of the theory.

VI. NUMERICAL RESULTS AND DISCUSSION

A. Chemical Terms

Table 2 lists $\delta_o = 1-d_I/d$, $\delta = 1-d_1/d$, the excess energy (per bond) ΔE for the full theory and its corresponding VFF Model D2, and the terms derived from the metallization energies: δ_o' (Eq.(19), F_{Ch}, H, ΔE_p, and ΔE_{Ch} (Eq.(17)]. The appropriate derivatives f, g, h, ..., [see Eq.(9)] are computed using the atomic term values that we have generated form impurity-level[13] and structural studies.[5]

For substitutions involving the cation pair (Ga, Al), F_{Ch} has the same sign as δ_o, which means that F_{Ch} prevents relaxation and thus tends to increase the strain energy. The chemical forces H are also significant. As a result, all six cases involving this pair have nearly equal d_1 and d_2, i.e. the small bond length differences are made even smaller in the alloy. The excess energies all become negative, mainly because ΔE_p is negative. For the systems involving the (Ga, In) and (In, Al) pairs, F_{Ch} has the opposite sign from δ_o, so δ_o' and δ_o have the same sign. The chemical force favors bond distortion. However, because H is positive and it introduces an increase in the denominator of Eq.(18), most of the effect of δ_o' is cancelled. For cases involving (Ga, In), the polarity contributions ΔE_p are all negative. The $F_{Ch}\Delta d$ term is negative, but $H\Delta d^2$ is positive, so they cancel to a certain degree and leave ΔE lowered primarily because of ΔE_p. While, ΔE_p is still negative for the (In,Al) substitutions, its magnitude is reduced considerably. The other chemical energies $F_{Ch}\Delta d + H\Delta d^2$ can be as

Table 2. Comparison between the full theory and the corresponding VFF Model D2 to study the effects of chemical terms. All ΔE's are in units of kcal/mol-band.

		Model D2			Full Theory						
	Host	δ_o	δ	ΔE	δ_o'	δ	$F_{ch}(10^{-10}N)$	H(N/m)	ΔE_p	ΔE_{ch}	ΔE
	AlP	0.003	0.002	0.001	-0.001	0.001	0.150	4.581	-0.016	-0.016	-0.013
l	GaP	-0.003	-0.002	0.001	0.001	-0.001	-0.150	4.581	-0.016	-0.016	-0.013
a	AlAs	0.001	0.001	0.000	-0.002	-0.001	0.243	5.733	-0.020	-0.021	-0.018
Al	GaAs	-0.001	-0.001	0.000	0.002	0.001	-0.243	5.733	-0.020	-0.021	-0.018
Ga	AlSb	0.006	0.004	0.005	-0.004	0.001	0.389	5.632	-0.054	-0.053	-0.039
Al	GaSb	-0.006	-0.004	0.005	0.004	-0.001	-0.389	5.632	-0.054	-0.053	-0.039
In	GAP	-0.077	-0.052	0.959	-0.006	-0.054	0.699	3.778	-0.188	-0.219	0.742
Ga	InP	0.071	0.056	0.734	0.005	0.057	-0.699	3.778	-0.188	-0.206	0.530
In	GaAs	-0.071	-0.048	0.752	-0.009	-0.050	0.804	4.778	-0.257	-0.283	0.472
Ga	InAs	0.066	0.052	0.592	0.007	0.054	-0.804	4.778	-0.257	-0.265	0.330
In	GaSb	-0.062	-0.043	0.554	-0.004	-0.042	0.352	5.201	-0.363	-0.308	0.247
Ga	InSb	0.059	0.046	0.445	0.004	0.044	-0.352	5.201	-0.363	-0.287	0.160
In	AlP	-0.074	-0.053	0.761	-0.007	-0.056	0.769	3.506	-0.035	-0.087	0.679
Al	InP	0.068	0.053	0.674	0.006	0.056	-0.769	3.506	-0.035	-0.083	0.596
In	AlAs	-0.070	-0.048	0.705	-0.010	-0.051	0.942	4.437	-0.048	-0.111	0.602
Al	InAs	0.065	0.052	0.576	0.008	0.054	-0.942	4.437	-0.048	-0.099	0.485
In	AlSb	-0.056	-0.039	0.440	-0.008	-0.041	0.689	4.979	-0.061	-0.073	0.369
Al	InSb	0.053	0.042	0.368	0.007	0.044	-0.689	4.979	-0.061	-0.061	0.310
Cd	ZnTe	-0.064	-0.048	0.432	-0.003	-0.050	0.202	-0.484	-0.005	-0.064	0.373
Zn	CdTe	0.060	0.050	0.314	0.002	-0.053	-0.202	-0.484	-0.005	-0.072	0.247
Hg	CdTe	0.003	0.002	0.001	0.004	0.005	-0.278	-0.753	-0.018	-0.026	-0.018
Cd	HgTe	-0.003	-0.002	0.001	-0.004	-0.005	0.278	-0.753	-0.018	-0.026	-0.018
Hg	ZnTe	-0.061	-0.045	0.392	-0.001	0.046	0.075	0.002	0.052	0.037	0.429
Zn	HgTe	0.058	0.048	0.286	0.001	0.049	-0.075	0.002	0.052	0.035	0.322
As	AlP	-0.03	-0.026	0.179	0.001	-0.025	-0.085	0.717	-0.005	0.008	0.187
P	AlAs	0.034	0.025	0.185	-0.001	0.025	0.085	0.717	-0.005	0.008	0.194
As	GAP	-0.037	-0.025	0.226	0.002	-0.024	-0.181	1.078	-0.011	0.012	0.240
P	GaAs	0.036	0.027	0.211	-0.001	0.025	0.181	1.078	-0.011	0.014	0.228
As	InP	-0.032	-0.023	0.136	0.001	-0.022	-0.057	0.919	-0.003	0.008	0.144
P	InAs	0.031	0.024	0.128	-0.001	0.024	0.057	0.919	-0.003	0.009	0.138
Sb	AlAs	-0.084	-0.058	1.024	0.008	-0.051	-0.815	0.644	-0.180	0.002	1.060
As	AlSb	0.077	0.059	0.919	-0.007	0.053	0.815	0.644	-0.180	0.027	0.984
Sb	GaAs	-0.078	-0.052	0.908	0.018	-0.040	-1.599	0.927	-0.363	-0.106	0.929
As	GaSb	0.073	0.055	0.823	-0.014	0.044	1.599	0.927	-0.363	-0.061	0.904
Sb	InAs	-0.070	-0.051	0.603	0.010	-0.042	-0.824	0.855	0.171	-0.009	0.645
As	InSs	0.065	0.051	0.551	-0.009	0.044	0.824	0.855	-0.171	0.008	0.613
Sb	AlP	-0.122	-0.085	2.007	0.010	-0.077	-0.944	0.645	0.241	0.074	2.127
P	AlSb	0.109	0.085	1.855	-0.008	0.078	0.944	0.645	-0.241	0.123	2.030
Sb	GAP	-0.119	-0.075	2.132	0.021	-0.061	-1.868	0.930	-0.505	-0.046	2.244
P	GaSb	0.106	0.083	1.806	-0.015	0.070	1.868	0.930	-0.505	0.093	2.084
Sb	InP	-0.104	-0.072	1.383	0.011	-0.063	-0.922	0.854	-0.214	0.059	1.501
P	InSb	0.094	0.077	1.191	-0.008	0.069	0.922	0.854	-0.214	0.123	1.379
Se	ZnS	-0.048	-0.036	0.231	-0.001	-0.036	0.077	0.645	0.003	0.000	0.231
S	ZnSe	0.046	0.037	0.221	0.001	0.037	-0.077	0.645	0.003	0.001	0.222
Te	ZnSe	-0.075	-0.056	0.550	0.000	-0.056	0.028	0.635	0.008	0.024	0.574
Se	ZnTe	0.069	0.054	0.532	0.000	0.054	-0.028	0.635	0.008	0.025	0.557
Te	ZnS	-0.126	-0.092	1.565	-0.001	-0.092	0.101	0.644	0.022	0.041	1.606
S	ZnTe	0.112	0.091	1.446	0.001	0.091	-0.101	0.644	0.022	0.051	1.496

large as ΔE_p. But the overall reductions of ΔE are only a fraction of those for the (Ga, In) cases. For the several II-VI systems studied, both F_{Ch} and H are small, and the net changes in δ have the same sign as δ_o. However, because δ_o is small in the (Cd, Hg) substitutions, F_{Ch} actually causes a reversal in which the short bond length gets shorter and the longer one gets longer. This is the only exceptional case of this type found for all the systems studied. The change of ΔE due to chemical terms in the (Hg, Zn) substitution is also peculiar--it increases mainly because ΔE_p is positive.

Next examine the anion substitutions. For the groups involving the (P, As) pair, the chemical shifts are all small, but the trend is less toward relaxation and larger ΔE. The groups involving (As, Sb) and (P, Sb) pairs behave very similarly: F_{Ch} are significant and are opposing relaxations, i.e. δ_o' and δ_o have opposite signs. At the same time, the H values are several times smaller than those for the corresponding III-V cation substitution case. Thus, most of δ_o' translates into a real reduction of the ratio δ/δ_o, and consequently introduces extra strain energy. Although the ΔE_p energies are significant and negative, $F_{Ch}\Delta d$ are positive, and the net ΔE_{Ch} can be either positive or negative. However, the induced strain energy due to reduction of the δ/δ_o makes all ΔE positive for these two groups of systems. For II-VI systems, all the chemical effects again are small, but the net chemical changes on ΔE are slightly positive.

The above discussion can be summarized in Figure 2, where the excess energies ΔE calculated from the full perturbation theory (FPT) and Model D2 are plotted normalized to the results of the simple spring model (SSHS)[8],of Shih et al. i.e. Eq.(24) with $\alpha_I = \alpha$, so $\Delta E/(3/8\ \alpha d^2) = \delta_o^2$. The calculated

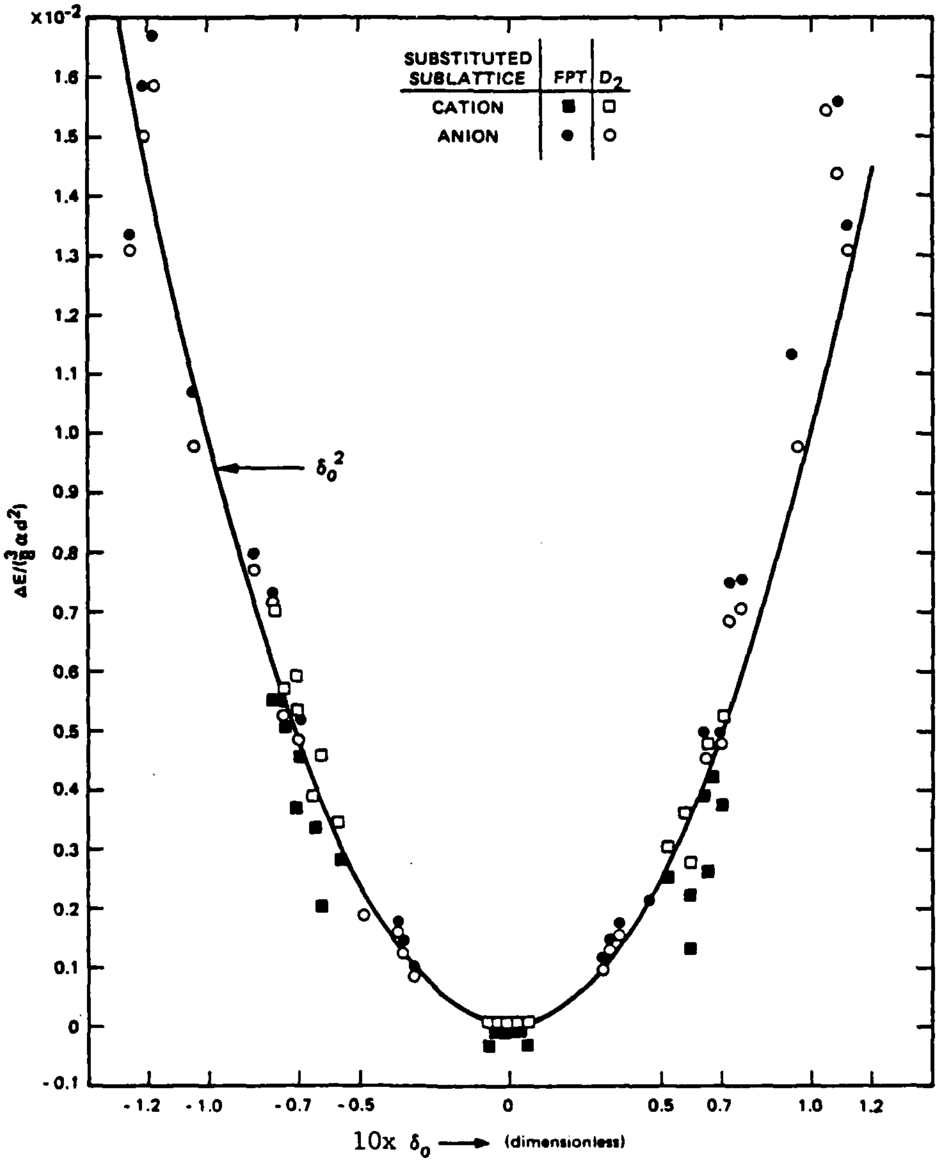


Fig. 2. The excess energies ΔE over $3/8\ \alpha d^2$ calculated from the full perturbation theory and its corresponding VFF Model D2. The solid curve corresponds to Eq. (24) with $\alpha = \alpha_I$.

ΔE rises faster for $\delta_o \geq o$ than for $\delta_o \leq o$, mainly because $\alpha/\alpha_I = 1$. In fact if the relation[7,10] $\alpha/\alpha_I = (d_I/d)^S$ with S of order of 3 to 5 is used in Eq.(24), we obtain a percentage correction of $S\delta_o/4$ to the SSHS results, which explains the skewed behavior of the curve. It is also clear that ΔE rises faster than δ_o^2 for larger δ_o. However, the zeroth order theory of SSHS is clearly an excellent representation of the global features of ΔE. The results from Model D2 are closer to the parobolic form than those from FPT. The figure clearly shows the general trends; the chemical terms cause negative shifts in ΔE for cation substitutions and positive shifts for anion impurities. It is also clear that the chemical shifts can be very large. These effects will have important consequences on the alloy mixing enthalpies to be discussed later.

B. Impurity Bond Length

Table 3 lists the impurity bond lengths d_1 calculated from different models, while a comparison of theory and the available experimental data[1,14,15] is presented in Table 4. The actual size of changes in d_1 induced by the chemical terms can be seen by comparing Model D2 with the full theory. Except for the systems involving the substitution of (Ga,As) and (P,Sb) pairs (where changes range from 0.01 to 0.03Å), all the chemically induced changes are less than 0.01 Å. Comparison among Models A, B, and C shows that, while extending the boundary helps the relaxation (compare Model B to A), i.e. δ/δ_o is closer to 1, the inclusion of the bond-bending forces (the β terms) (compare Model B with C) prevents it. The simple spring model (Model C) which contains neither of these

Table 3. Calculated impurity local bond lengths (in Å) from the full theory and several valence force models discussed in Section III and their comparison with the values calculated by Martins and Zunger (Ref. 3)

		A	B	C	D1	D2	E	FPT	MZ
Ga	AlP	2.362	2.363	2.362	2.362	2.362	2.364	2.364	
Al	GaP	2.365	2.364	2.365	2.365	2.365	2.365	2.363	
Ga	AlAs	2.449	2.449	2.449	2.449	2.449	2.450	2.452	
Al	GaAs	2.450	2.450	2.450	2.450	2.450	2.450	2.447	
Ga	AlSb	2.645	2.646	2.644	2.645	2.644	2.649	2.653	
Al	GaSb	2.651	2.650	2.652	2.652	2.652	2.650	2.644	
In	GAP	2.477	2.462	2.492	2.479	2.483	2.435	2.487	2.474
Ga	InP	2.406	2.421	2.402	2.395	2.399	2.409	2.396	2.409
In	GaAs	2.559	2.544	2.573	2.561	2.565	2.518	2.570	2.556
Ga	InAs	2.492	2.506	2.486	2.482	2.485	2.496	2.481	2.495
In	GaSb	2.747	2.734	2.760	2.749	2.754	2.710	2.750	2.739
Ga	InSb	2.683	2.697	2.678	2.673	2.676	2.686	2.680	2.683
In	AlP	2.487	2.472	2.494	2.490	2.493	2.447	2.498	2.480
Al	InP	2.412	2.427	2.408	2.401	2.405	2.415	2.400	2.414
In	AlAs	2.561	2.546	2.572	2.563	2.568	2.523	2.575	2.553
Al	InAs	2.493	2.506	2.487	2.483	2.487	2.497	2.480	2.495
In	AlSb	2.754	2.741	2.763	2.756	2.760	2.721	2.765	2.746
Al	InSb	2.693	2.705	2.689	2.685	2.687	2.696	2.683	2.693
Cd	ZnTe	2.756	2.740	2.760	2.760	2.763	2.720	2.770	2.755
Zn	CdTe	2.673	2.688	2.676	2.660	2.665	2.671	2.658	2.674
Hg	CdTe	2.800	2.801	2.800	2.800	2.799	2.801	2.790	
Cd	HgTe	2.804	2.803	2.804	2.805	2.805	2.804	2.813	
Hg	ZnTe	2.750	2.735	2.754	2.753	2.757	2.715	2.758	2.748
Zn	HgTe	2.671	2.685	2.674	2.659	2.664	2.671	2.662	2.673
As	AlP	2.425	2.418	2.429	2.427	2.428	2.406	2.427	2.422
P	AlAs	2.392	2.399	2.387	2.387	2.389	2.394	2.390	2.395
As	GAP	2.417	2.409	2.424	2.417	2.420	2.396	2.416	2.414
P	GaAs	2.386	2.393	2.380	2.382	2.383	2.389	2.386	2.387
As	InP	2.596	2.589	2.599	2.598	2.600	2.579	2.598	2.595
P	InAs	2.561	2.568	2.558	2.557	2.558	2.563	2.560	2.562
Sb	AlAs	2.584	2.566	2.597	2.587	2.592	2.539	2.577	2.574
As	AlSb	2.506	2.522	2.496	2.495	2.498	2.511	2.514	2.510
Sb	GaAs	2.569	2.553	2.584	2.571	2.576	2.524	2.546	2.564
As	GaSb	2.501	2.516	2.489	2.492	2.495	2.508	2.525	2.505
Sb	InAs	2.747	2.730	2.754	2.750	2.754	2.705	2.733	2.739
As	InSb	2.669	2.683	2.663	2.658	2.662	2.672	2.683	2.667
Sb	AlP	2.555	2.529	2.569	2.561	2.568	2.488	2.550	2.542
P	AlSb	2.440	2.462	2.426	2.425	2.430	2.447	2.448	2.444
Sb	GAP	2.526	2.503	2.551	2.529	2.537	2.461	2.504	2.519
P	GaSb	2.431	2.451	2.414	2.418	2.422	2.440	2.454	2.436
Sb	InP	2.712	2.687	2.720	2.719	2.725	2.654	2.702	2.700
P	InSb	2.599	2.619	2.591	2.585	2.590	2.604	2.611	2.597
Se	ZnS	2.420	2.409	2.421	2.424	2.426	2.396	2.426	2.420
S	ZnSe	2.367	2.376	2.365	2.360	2.364	2.370	2.363	2.367
Te	ZnSe	2.586	2.569	2.588	2.589	2.592	2.540	2.591	2.584
Se	ZnTe	2.501	2.517	2.497	2.490	2.494	2.504	2.495	2.502
Te	ZnS	2.543	2.513	2.544	2.552	2.558	2.478	2.557	2.539
S	ZnTe	2.406	2.429	2.400	2.390	2.396	2.410	2.396	2.407

Table 4. Deviations of the Calculated Impurity Bond Lengths (INÅ) as Compared with the Experimental Values from EXAFS

Impurity	Host	A	C	D1	FPT	MZ	SSHS[*]	Exp.
In	GaAs	-0.028	-0.014	-0.025	-0.017	-0.031	-0.009	2.587[†]
Ga	InAs	0.004	-0.001	-0.005	-0.006	0.008	0.005	2.487[†]
Cd	ZnTe	0.004	0.008	-0.008	0.018	0.003	0.012	2.75(2)[‡]
Zn	CdTe	-0.01 ~ -0.02	0. ~ -0.01	-0.02 ~ -0.03	-0.02 ~ -0.03	-0.01 ~ -0.02	0.0 ~ -0.01	2.68 ~ 2.69[‡]
Te	ZnSe	-0–009	-0.07	-0.003	-0.004	-0.011	-0.004	2.595[§]
Se	ZnTe	0.005	0.01	-0.006	-0.001	0.006	0.004	2.496[§]
Mean Absolute Deviation		0.011	0.006	0.012	0.012	0.012	0.006	

---

[*] SSHS: Ref. 8

[†] Ref. 1

[‡] Quoted in Ref. 3

[§] Ref. 14

terms, evidently represents a delicate cancellation of these effects and predicts results close to those of the full perturbation theory and experiment. Although the d_1 values of Model C are often very close to those of Model A, there are cases [e.g., Ga(P, Sb)] in which Model C can differ from Model A by 0.025 A. Model A produces about the same d_1 values as Model D1, where the maximum difference in d_1 is only 0.015 A. Martins and Zunger[3] used the same model as Model A, however, their analytic expression for δ is different from that given by Eq.(21). Nevertheless numerical results indicate these two calculations agree to 0.01 A. The slightly different forms of strain energies used in Model D2 and D1 only introduce a small change in d_1 with the largest difference being less than 0.01 A. The first-shell continuum model (Model E) allows too little relaxation, so while the other models produce a ratio δ/δ_o ranging from 0.6 to 0.8, Model E only ranges from 0.4 to 0.6. The reason that the fixed boundary in Model A works is that the effective shear coefficient C (see Table 1) characterizing the strain energy in the elastic continuum is large. However, Model B is too rigid, and does not provide enough buffer between the impurity bond and the fixed boundary.

The comparison of the theoretical results with the available experimental data in Table 4 indicates that Models B and E are the least accurate. Models A, D1, MZ and the full theory are comparable in that all have an average absolute deviation of 0.012 A which is about the experimental uncertainty in EXAFS. The agreement between theory and experiment, however, is not uniform. The most surprising result in Table 4 is that the simple spring model, Model C, and its cruder version used by Shih et al.[8] ($\alpha = \alpha_I$ so $\delta/\delta_o = 0.75$, labeled as SSHS) has the smallest variance in d_1, about 0.006 A. We know this does not imply that the simple spring model represents the real picture of bond-length relaxation. Nature is just playing tricks on us again. For example, if we let all the shear coefficients be zero, i.e., $\beta = C = 0$ in our model, then as the range of the boundary is gradually extended, the local bond length will eventually relax to the impurity bond length $d_1 = d_I$, or $\delta = \delta_o$. This can be seen in Model A from Eq. (21), where δ reduces to $\delta_o/(1+\alpha/6\alpha_I)$, rather than $\delta_o/(1+\alpha/3\alpha_I)$ as predicted by Model C, and in Model D from Eqs.(27) and (30), δ becomes δ_o, if the continuum is taken to be shearless. Considering that various effects are included that may mask the absolute accuracy of d_1 predictions (e.g. while low-temperature bond lengths are used in the calculation, the room-temperature values of elastic constants are adopted), the agreement of various models with experiments in Table 4 should be regarded as excellent. There are, however, many other impurity systems in which the simple spring model predictions differ considerably from other models, as is shown in Figure 3, where δ is plotted against δ_o for the full theory. Those points that deviate significantly from the 0.75-slope line are the systems with (As, Sb) or (P, Sb) substitutions. Additional EXFAS measurements on these systems are needed to test these predictions.

C. Mixing Enthalpies

Table 5 lists the mixing enthalpy parameters Ω (in kcal/mole) for a number of alloys estimated from Eq.(41) for all the models considered along with other theoretical[3,16-18] and experimental values.[19] As already discussed, the chemical terms reduce the excess energies in the cation

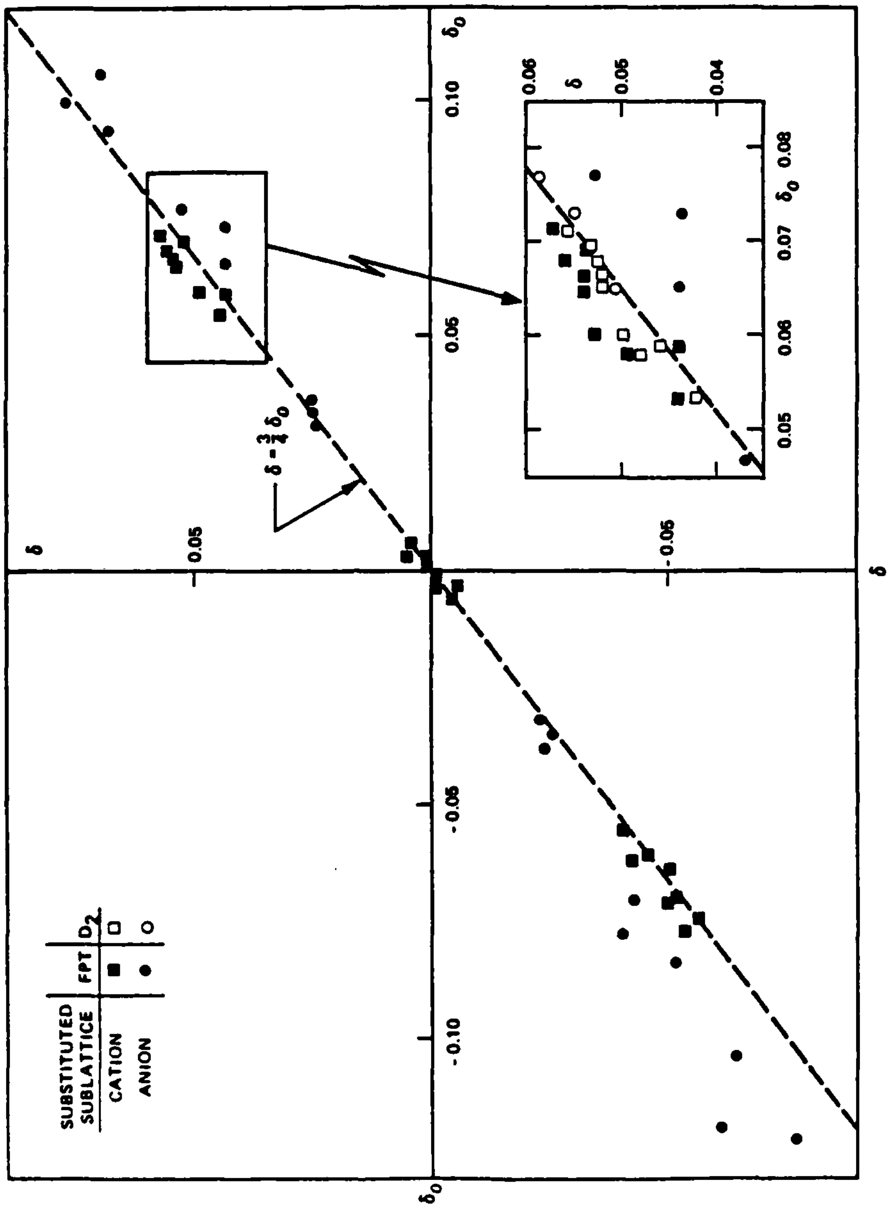


Fig. 3. Calculated bond relaxation parameter δ from FPT and D2 as a function of δ_0. The $\delta = 3/4\ \delta_0$ curve corresponds to theory of SSHS (Ref. 8).

Table 5. Mixing enthalpy parameter Ω (in kcal/mole) estimated from the full perturbation theory (FPT) and several valence force models discussed in Section III and comparison with experiments and other theories.

	A	B	C	D1	D2	E	FPT	MZ[a]	DL[b]	FM[c]	VV[d]	Exp.[e]
(Ga,Al)P	0.00	0.01	0.00	0.01	0.01	0.01	-0.05					
(Ga,Al)As	0.00	0.00	0.00	0.00	0.00	0.00	-0.07	0.02	0.02	0.03	0.11	0.0
(Ga,Al)Sb	0.02	0.03	0.02	0.02	0.02	0.03	-0.15	0.02	0.02	0.03		0.0
(Ga,In)P	3.76	4.79	3.0.	3.29	3.39	5.24	2.54	4.56	3.63	2.94		3.25 3.5
(Ga,In)As	2.97	3.76	2.36	2.61	2.69	4.14	1.60	2.49	2.81	2.42	1.25	1.65 2.0 3.0
(Ga,In)Sb	2.22	2.83	1.77	1.95	2.00	3.09	0.81	2.53	1.85	1.83		1.47 1.9
(In,Al)P	3.24	4.22	2.77	2.78	2.87	4.60	2.55					
(In,Al)As	2.86	3.65	2.32	2.49	2.56	3.93	2.17	3.60	2.81	2.37		2.5
(In,Al)Sb	1.81	2.33	1.49	1.57	1.61	2.50	1.36	2.06	1.46	1.45		0.6
(Cd,Zn)Te	1.80	2.43	1.73	1.43	1.49	2.45	1.24	2.12	1.97	1.63		1.34
(Hg,Cd)Te	0.00	0.00	0.00	0.00	0.00	0.00	-0.07					0.7 1.4
(Hg,Zn)Te	1.63	2.20	1.56	1.30	1.36	2.23	1.50	1.91	1.81	1.48		3.0
Al(P,As)	0.81	1.03	0.65	0.71	0.73	1.14	0.76					
Ga(P,As)	0.95	1.18	0.70	0.86	0.87	1.32	0.94	1.15	0.98	0.66	0.12	0.4 1.0
In(P,As)	0.60	0.78	0.52	0.51	0.53	0.84	0.57	0.72	0.58	0.52		0.4
Al(As,Sb)	4.31	5.45	3.38	3.80	3.88	5.92	4.09					
Ga(As,Sb)	3.77	4.69	2.81	3.40	3.46	5.22	3.67	4.58	3.35	2.76		4.0 4.5
In(As,Sb)	2.61	3.39	2.23	2.24	2.31	3.67	2.52	2.89	2.29	2.17	6.65	2.25 2.9
Al(P,Sb)	8.60	10.99	6.99	7.54	7.73	12.00	8.32					
Ga(P,Sb)	8.54	10.61	6.36	7.72	7.88	11.66	8.66					
In(P,Sb)	5.87	7.64	5.08	4.99	5.15	8.04	5.76					
Zn(S,Se)	1.04	1.39	0.98	0.85	0.90	1.49	0.90					
Zn(S,Se)	1.04	1.39	0.98	0.85	0.90	1.49	0.90					
Zn(Se,Te)	2.47	3.27	2.23	2.09	2.16	3.63	2.26	2.91	3.11	2.12	3.12	1.55
Zn(S,Te)	7.02	9.34	6.45	5.80	6.02	9.72	6.20					

---

[a] Ref. 3, Col. A of Table III.

[b] Ref. 16

[c] Ref. 13

[d] Ref. 17

[e] Ref. 16 and 19.

impurities and increase them for anion impurities. The corresponding changes in Ω are the differences between the FPT and Model D2. We note that the reductions of Ω for the (Ga, In) alloys are very large (> 1 kcal/mole) and also significant for (In, Al) alloys. However, the increases in Ω for the anion substitutional alloys are not as large. Also the Ω differences between Models D1 and D2 are less than 10 percent. Model A produces Ω values about 20 percent larger than Model D1; Model B in turn is 20 percent higher than Model A; and Model E is 10 percent higher than Model B. The Ω values in the simple spring model, Model C, are seen to be about the same as Model D1, although the differences among systems can be positive or negative. Although MZ used the same strain model as Model A, their Ω values do not agree with our Model A values because their way of estimating Ω is different. In fact, MZ's values are closer to Model B than A.

To distinguish the quantitative nature of different theoretical models, we note that there are also important factors that may mask the comparison between theory and experiment for Ω. One important factor is that the mixing enthalpies extracted from phase diagram analysis are sensitive to sample and experimental conditions. These ΔH_m contain contributions from various nonideal structures such as vacancies, impurities, dislocations, grain boundaries, and surface conditions, in addition to the ideal ΔH_m considered here for solid solutions. Thus, our theoretical ΔH_m should represent a lower bound. Another complication comes from the version of the theory of solid solution adopted. The theory used for analysis so far assumes a regular solid solution with second-neighbor pair interactions as was outlined in Section V. Recent experimental[20-22] and theoretical[23] studies have suggested the possibility of compositional clustering or long-range correlations in alloys. Extending the theory to include such effects will alter the simple results for ΔH_m from Eq.(41). Moreover, there is evidence from the composition variation of the alloy hardness[24] and from the optical phonon frequencies [25] that the shear coefficients of alloys increase near the center of the composition range. This will cause the effective continuum shear coefficient C in Eq.(14) to be composition dependent, which will cause Ω to increase. Despite these uncertainties, useful comparisons across the board in Table 5 can still be made.

Based on the above considerations, we can conclude that Models B, E, and MZ predict Ω values that are too high. We should emphasize that all the Ω numbers for models from A through MZ are directly calculated without any adjustable parameters. The fact that Models A, C, and D1, D2, and the FPT agree with the experiment as well as or even slightly better than the one-parameter theories[16,18] (DL and FM) is already quite an accomplishment. The few numbers taken from Van Vechtan's calculations[17](VV) indicate that the dielectric model predicts results at larger varience with experiments. There are two important implications about the FPT that can be drawn from Table 3. First, the theory predicts a small but negative Ω value for several alloys. This not only means that there is no miscibility gap in these alloys but also implies a tendency toward ordering, in which the substitutional atoms tend to be surrounded by different second-neighbor species. For stoichiometric compositions, this implies a tendency toward compound formation. Secondly, the FPT tends to predict smaller Ω values than observed experimentally, which should be expected according to our discussion. To the extent that the FPT predicts the correct ΔH_m values for an ideal solution, the difference $\Delta H_m^{exp} - \Delta H_m$ may be attributed to imperfect conditions and deviations from the ideal solution theory.

Finally, the calculated Ω values in Table 5 provide some guidance in separating the completely miscible alloys from immiscible ones.[26,27] In a true random alloy, the criteria[28] for alloy mixing at a temperature T is that $T \geq T_c$, where the critical temperature T_c is given by $\Omega/(2R_g)$ with R_g being the universal gas constant.[29] For an $A_xB_{1-x}C$ alloy to be miscible throughout the whole concentration range, the requirement is that both the melting temperatures T_1 and T_2 of the pure AC and BC compounds are greater than T_c. Table 6 lists the values of T_c associated with the Ω values in the FPT, the ratios T_c/T_1 and T_c/T_2, and the average absolute values of δ_o for the alloys considered in Table 5. In Table 6, T_c is set equal to zero if Ω is negative and T_2 is chosen to be the lower value of the two melting temperatures, so the criterion for not having a miscibility gap is $T_c/T_2 < 1$. There is an empirical rule[26] stating that this will be satisfied if the lattice mismatch $|\delta_o|$ between the two alloy components is less than 7.5 percent. However, we find that (see Appendix C) a more precise rule is that $|\delta_o| < \delta_m$, where $\delta_m = 1.63\ \chi_m$ and χ_m is the ratio of the rms bond length amplitude fluctuation to the bond length at the melting temperature T_2. The values of T_m for the compounds involved[30] and the associate χ_m values estimated from Eqs.(C2) and (C3) are tabulated in Table 1. The model used in Appendix C yields $T_c/T_2 = (\delta_o/\delta_m)^2$. This suggests that it is instructive to plot T_c/T_2 as a function of $|\delta_o|/\delta_m$, as is shown in Figure 4 for the T_c calculated from FPT. This plot is similar to the ΔE vs δ_o curve in Fig. 2, because, in fact, Ω is proportional to the sum of the ΔE values of the two constituents [see Eq. (41)]. However, if T_c/T_2 is plotted against $|\delta_o|$

Table 6. Comparison of the Critical Temperature T_c of Mixing and Melting Temperatures of the Constituents T_1 and T_2, in the Order of Their Appearance in the Parentheses. Also Tabulated are the Averaged Absolute Values of $|\delta_o|$ and the Ratio $|\delta_o|/\delta_m$

System	$|\delta_o|$ (%)	T_c (°K)	T_c/T_1	T_c/T_2	$|\delta_o|/\delta_m$
(Al,Ga)As	0.1	0	0	0	0.009
(Al,Ga)P	0.3	0	0	0	0.029
(Al,Ga)Sb	0.6	0	0	0	0.067
(Al,In)Sb	5.5	342	0.25	0.42	0.679
(Ga,In)Sb	5.8	204	0.21	0.25	0.716
(Al,In)As	6.8	547	0.29	0.45	0.687
(Ga,In)As	6.9	403	0.23	0.34	0.697
(Al,In)P	7.1	642	0.36	0.48	0.732
(Ga,In)P	7.4	639	0.43	0.48	0.763
In(P,As)	3.2	144	0.11	0.12	0.330
Al(As,P)	3.5	191	0.10	0.11	0.307
Ga(As,P)	3.7	236	0.14	0.15	0.352
In(As,Sb)	6.8	635	0.52	0.79	0.840
Ga(As,Sb)	7.6	924	0.53	0.94	0.844
Al(As,Sb)	8.1	1030	0.56	0.78	0.810
In(P,Sb)	9.9	1450	1.08	1.82	1.222
Ga(P,Sb)	11.3	2180	1.25	2.21	1.256
Al(P,Sb)	11.6	2095	1.19	1.58	1.116
(Cd,Hg)Te	0.3	0	0	0	0.033
(Zn,Hg)Te	6.0	377	0.25	0.40	0.659
(Zn,Cd)Te	6.2	312	0.21	0.23	0.564
Zn(S,Se)	4.7	226	0.11	0.13	0.362
Zn(Se,Te)	7.2	569	0.32	0.38	0.615
Zn(S,Te)	11.9	1560	0.74	1.03	1.017

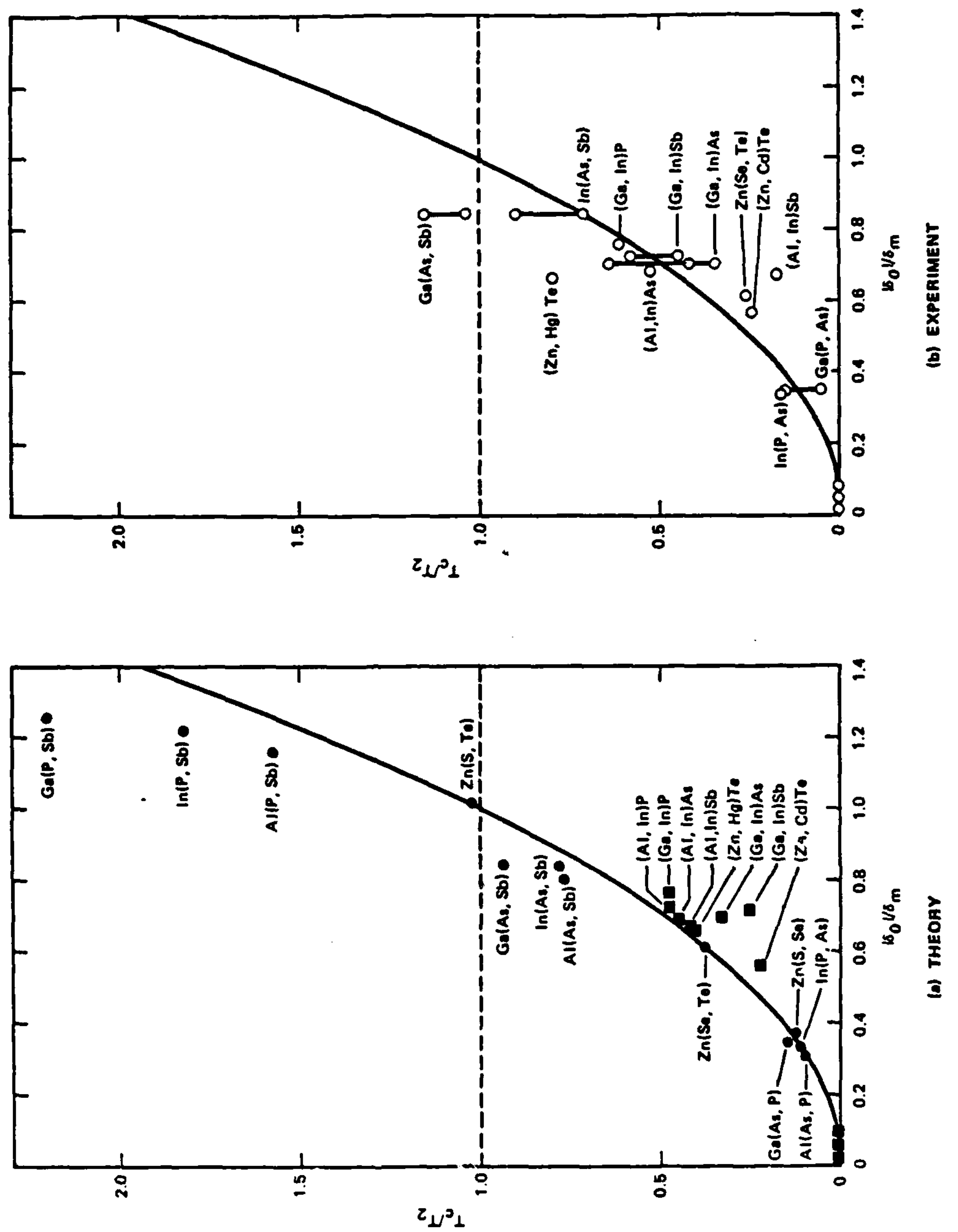


Fig. 4. Plot of T_c/T_2 as a function $|\delta_o|/\delta_m$ for T_c obtained from (a) FPT, and (b) the experimental Ω in Table 5. In (a), the dots are for anion alloys and crosses for cation substitution. The solid lines in both parts correspond to the simple theory discussed in Appendix C. The dashed lines at $T_c/T_2 = 0$ separate the miscible from immiscible groups.

alone, the FPT points are much more scattered, and those of SSHS would not even exhibit a smooth simple quadratic form because the lower melting temperature T_2 is not a smooth function of $|\delta_o|$. This result suggests that $|\delta_o|/\delta_m < 1$ is a better criterion than $|\delta_o| < 0.075$. Figure 3 also clearly shows the chemical effects: all the cation-substitution alloy points lie below the solid curve and all the anion-substitution alloys have (T_c/T_2) values on or above the curve, corresponding to negative and positive shifts in ΔE due to the chemical terms. Again, the curve based the SSHS model is an excellent universal representation. From the figure, we see that all (P,Sb) alloys should have miscibility gaps and all (As, Sb) alloys are predicted to be miscible, although on the borderline, because the actual mixing enthalpies are larger than these ideal calculated values. The figure also shows that Zn(S, Te) has a miscibility gap but a smaller value of T_c/T_2 than the (P, Sb) alloys, despite the fact that its $|\delta_o|$ value is larger. All these predictions are consistent with the available experimental evidence.

VII. SUMMARY AND CONCLUSION

In this paper, a simple theory of defect substitution energies is formulated. The substitution energy is compactly separated into a replacement energy Δ_R and a distortion energy of the pure host crystal [see Eq.(2)]. However, a rigorous application of this theory requires an improvement in certain quantitative aspects of Harrison's bonding theory,[7] particularly the elastic constants.[5] The most interesting application of this theory presented in this paper is its perturbation form which enables us to absorb the measured elastic coefficients into the calculation, and more importantly, to study the chemical effects. The origin of chemical influences on impurity bond relaxation can be attributed to three mechanisms [see Eq.(17)]: a chemical force F_{Ch} that either helps or hinders lattice relaxation, depending on whether it has the same or opposite sign from the bond length difference $d-d_I$ between the host and impurity, a chemical energy that depends on the difference of the polarities between the impurity and host bonds, ΔV_3, and an effective elastic force constant H that, when positive, also tends to restrain the lattice from distortion. To study the effect of boundaries between the core atoms around an impurity and the rest of the elastic medium, various models based on the valence force field[10] are derived and their results are compared with the full perturbation theory and available exerimental data. We found at least five models, including the FPT, that produce the correct impurity bond lengths with variances for the compounds studied about equal to the experimental uncertainties in EXAFS[1,4,14] (~ 0.01 Å). However, some models are oversimplified and will certainly not predict other properties equally well. However, more experimental lattice constant measurements, to further test the theory, particularly on (As, Sb) and (P,Sb) substitution systems for which there are larger differences between different models, are needed. It would also be instructive to see if the predicted reversal for HgCdTe is found.

The excess energies of impurity substitution are also shown to provide good estimates of the mixing enthalpies Ω of pseudo-binary alloys. The chemical shifts are found to have a negative net contribution to Ω for most cation substitutions, but positive contributions for anion substitutions. The chemical reduction of Ω in (Ga, In) alloys is larger than 1 kcal/mole (30 percent to 100 percent). Several VFF models and the full perturbation theory produce results for Ω that are as good as the best theories with one adjustable parameter. However, the full theory tends to yield answers on the low side of the experimental values, which we argue is as it should be, because there are nonideal structures that also contribute to Ω. The calculated Ω values and the melting temperatures are used to predict the existance of alloy miscibility gaps, and the results correlate well with experiments.

Finally, we wish to comment on the accuracy of the theories that are connected to the present model. The perturbation theory has already been stretched beyond its expected region of validity and predicts d_1 to within experimental uncertainties ($\sim$ 0.01 Å) even for cases with large bond length differences ($\delta_0 \sim 0.1$). The accuracy can only be improved if the full non-perturbation theory outlined in Section II is used. This calculation is needed for the strong substitution cases that were not considered in this paper: examples are (B, Ga), (B, In), (N,P), (N,As), (N,Sb), (O,S), (O,Se) and (O,Te) substitutions. Although we believe that, for the properties treated, the model with a continuum attached to the second-shell is as accurate as the perturbation theory used, it remains to be seen if this is true for other properties, especially strain coefficients. Finally, the present theory has been extended to study alloys[5,6] by embedding clusters in an effective medium. This enables us to study the bond length and energy variations throughout the whole concentration range. However, a quantitative calculation still awaits an improvement of the accuracy of Harrison's theory. A similar precedure is also being extended to a study of the alloy electronic structure, for which a cluster CPA (coherent-potential approximation) involving both potential and structural disorder[31] will be used.

ACKNOWLEDGEMENT

This work benefited from useful discussions with Professors W. A. Harrison and A. Zunger. A.-B Chen would like to thank Professor W. E. Spicer for his hospitality. The work is supported in part by AFOSR-84-0284 and DARPA Contract MDA 903-83-C-0108.

Appendix A

ELASTIC ENERGY IN CONTINUUM

In Section 3, the elastic energy outside a sphere of radius R centered at the impurity is assumed to be a continuum with a radial displacement $\vec{u}\alpha\hat{r}/r^2$. If the displacement at R is u_o, then $\vec{u}(\vec{r}) = u_o (R^2/r^2)\hat{r}$. The energy density in the continuum is given by

$$\delta\varepsilon(\vec{r}) = \frac{1}{2} C_{11}(e_{xx}^2+e_{yy}^2+e_{zz}^2) + C_{12}(e_{yy}e_{zz}+e_{xx}e_{yy}+e_{xx}e_{zz})$$

$$+ \frac{1}{2}C_{44}(e_{xy}^2+e_{yz}^2+e_{zx}^2),$$

where

$$e_{xx} = \frac{\partial u_x}{\partial x} = R^2u_o(r^2-3x^2)/r^5,$$

$$e_{xy} = \frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial x} = -6R^2u_o xy/r^5, \ \ldots$$

Thus the total elastic energy in the continuum is

$$\Delta_{\text{dis}}^{(\text{out})} = \int_R^\infty \delta\varepsilon(\vec{r})\, d^3r$$

$$= 4\pi R u_o^2 (\frac{2}{5}\, C_{11} - \frac{2}{5}\, C_{12} + \frac{6}{5} C_{44})$$

$$= C R u_o^{\,2},$$

where *the effective shear coefficient is given by*

$$C = \pi(1.6C_{11} - 1.6C_{12} + 4.8C_{44}).$$

Appendix B

DISTORTION ENERGY

In this appendix, we count the detailed contributions of the bond stretching terms $\Delta(\vec{r}_i \cdot \vec{r}_i)$, and "bond bending" terms $\Delta\ (\vec{r}_i \cdot \vec{r}_j)$ for $i \neq j$ in VFF [Eq.(13)] that enter Eq.(16) in FPT and in the VFF Models in Section III.

α and β terms from the first-shell bonds

The four bond vectors pointing away from the central impurity according to Fig. 1 are $\vec{r}_1 = (1-\delta, 1-\delta, 1-\delta)d/\sqrt{3}$. $\vec{r}_2 = (1-\delta, -1-\delta, -1-\delta)d/\sqrt{3}$ ... Thus $\Delta(\vec{r}_1 \cdot \vec{r}_1) = -2\delta d^2$ and $\Delta(\vec{r}_1 \cdot \vec{r}_2) = -\frac{2}{3}\delta d^2$. The α-terms contribute $4 \cdot 3\alpha_I$ $(-2\delta d^2)^2/8d^2 = 6\alpha_I \delta^2 d^2$, and the β-terms contribute $6 \cdot 3\beta_I(-\frac{2}{3}\ \delta d)^2/8d^2 = \beta_I \delta^2 d^2$. If an A atom is replaced by a B atom, as was done in FPT, the α_I and β_I are replaced by α and β respectively.

α-terms from the 2nd-shell bonds, β-terms between the first- and second-shell bonds and among the 2nd-shell bonds

For these terms we need to consider the four bond vectors pointing away from C. They are $\vec{r}_1 = (-1+\delta, -1+\delta, -1+\delta)d/\sqrt{3}$, $\vec{r}_2 = (-1+\delta,\ 1+\delta+\gamma,\ 1+\delta+\gamma)\ d/\sqrt{3}$, $\vec{r}_3 = (1+\delta+\gamma,\ -1+\delta,\ 1+\gamma+\delta)d/\sqrt{3}$, ... Then $\Delta\ (\vec{r}_2 \cdot \vec{r}_2) = \frac{2}{3}\ (\delta+2\gamma)d^2$,

$\Delta(\vec{r}_1 \cdot \vec{r}_2) = -\frac{2}{3}(\delta+\gamma)d^2$, and $\Delta(\vec{r}_2 \cdot \vec{r}_3) = \frac{2}{3}\delta d^2$. Thus the α-terms from the second-shell bonds become $4 \cdot 3 \cdot 3\alpha\ [\frac{2}{3}(\delta+2\gamma)\ d^2]^2/8d^2 = 2\alpha\ (\delta+2\delta)^2 d^2$, the β-term between the first- and second-shell bonds are $4 \cdot 3 \cdot 3\beta\ [\frac{2}{3}(\delta+\gamma)d^2]^2/8d^2 = 2(\gamma+\delta)^2\beta d^2$, and the β-terms among the second-shell bonds are $4 \cdot 3 \cdot 3\beta$ $(\frac{2}{3}\ \delta d^2)^2/8d^2 = 2\delta^2\beta d^2$.

α-terms for the third-shell bonds, β-terms between the second- and third-shell bonds and among the third-shell bonds adjacent to the second-shell atoms

For these terms we need to consider the bond vectors pointing away from B in Figure 1. They are $\vec{r}_2 = (1-\delta, -1-\gamma-\delta, -1-\gamma-\delta)d/\sqrt{3}$, $\vec{r}_1 = (1+\gamma', 1+3\gamma'-\gamma, 1+3\gamma'-\gamma)d/\sqrt{3}$, $\vec{r}_3 = (-1-\gamma'', 1+3\gamma''-\gamma, -1+\gamma''-\gamma)d/\sqrt{3}$ and $\vec{r}_4 = (-1-\gamma'', -1+\gamma''-\gamma, 1+3\gamma''-\gamma)\ d/\sqrt{3}$. Thus we have $\Delta\ (\vec{r}_1 \cdot \vec{r}_2) = -\frac{d^2}{3}\ (3\delta+5\gamma')$, $\Delta(\vec{r}_2 \cdot \vec{r}_3) = \frac{d^2}{3}\ (\delta+2\gamma-5\gamma'')$, $\Delta(\vec{r}_3 \cdot \vec{r}_4) = -\frac{2}{3}\ d^2\gamma''$, $\Delta(\vec{r}_1 \cdot \vec{r}_3) = \frac{d^2}{3}\ (-2\gamma-\gamma'+3\gamma'')$, $\Delta(\vec{r}_3 \cdot \vec{r}_3) = \frac{d^2}{3}\ (6\gamma'')$, and $\Delta(\vec{r}_1 \cdot \vec{r}_1) = \frac{d^2}{3}\ (14\gamma'-4\gamma)$. For Model A, $\gamma' = r'' = 0$, so the α terms of the third shell bonds become $4 \cdot 3 \cdot 3\alpha\ [2(\Delta\vec{r}_3 \cdot \vec{r}_3)^2 + (\Delta\vec{r}_1 \cdot \vec{r}_1)^2]/8d^2 = 8\alpha\gamma^2 d^2$, the β terms between the 2nd- and 3rd-shell bonds are $4 \cdot 3 \cdot 3\beta[(\Delta\ \vec{r}_1 \cdot \vec{r}_2)^2 + 2(\Delta\ \vec{r}_2 \cdot \vec{r}_3)^2]/8d^2 = \frac{9}{2}\ \beta d^2[\delta^2 + \frac{2}{9}\ (\delta+2\gamma)^2] = \frac{9}{2}$ $\beta\delta^2 d^2 + \beta(\delta+2\gamma)^2 d^2$, and the β terms among the third-shell bonds adjacent to the second-shell atoms are $4 \cdot 3 \cdot 3\ \beta[2(\Delta\vec{r}_1 \cdot \vec{r}_3)^2/(8d^2) + (\Delta\vec{r}_3 \cdot \vec{r}_4)^2] = 4\beta\gamma^2 d^2$. For continuumm and the only contribution from this group are the β terms between the second- and third-shell bonds. Since the displacements in the continuum are proportional to $1/R^2$, $\gamma' = 8\sqrt{2}\,\gamma/(19\sqrt{19})$ and $\gamma'' = 8\sqrt{2}\gamma/(11\sqrt{11})$. Thus, these β terms become

$$\frac{9}{2d^2}\,\beta\left[\frac{d^4}{9}(3\delta+5\gamma')^2+\frac{2d^4}{9}\,(\delta+2\gamma-5\gamma'')^2\right] = \frac{1}{2}\,\beta\left[\left(3\delta+\frac{40\sqrt{2}}{19\sqrt{19}}\gamma\right)^2 + 2\left(\delta+2\gamma-\frac{40\sqrt{2}}{11\sqrt{11}}\gamma\right)^2\right]d^2.$$

β terms for bonds adjacent to the third-shell atoms

These terms only enter Model A, so r' = r" = 0. There are two different groups, one like those adjacent to C' and another like those meeting at C". The four bond vectors pointing away from C' are $\vec{r}_1 = (-1, -1+\gamma, -1+\gamma)\ d/\sqrt{3}$, $\vec{r}_2 = (-1,1,1)\ d/\sqrt{3}$, $\vec{r}_3 = (1,-1,1)\ d/\sqrt{3}$, and $\vec{r}_4 = (1,1,-1)\ d/\sqrt{3}$. Thus the only contribution from this group is $4\cdot3\cdot3\beta\,[\Delta(\vec{r}\cdot\vec{r}_2)]^2/8d^2 = 2\beta\gamma^2d^2$. The four bond vectors around C" are: $\vec{r}_1 = (-1+\gamma, -1-\gamma, -1)d/\sqrt{3}$, $\vec{r}_2 = (-1,1,1)d/\sqrt{3}$, $r_3 = (1,-1+\gamma, 1+\gamma)d/\sqrt{3}$, and $\vec{r}_4 = (1,1,-1)d/\sqrt{3}$, which only results in the first order term $\Delta(\vec{r}_2\cdot\vec{r}_3) = \frac{2}{3}\,\gamma d^2$. Thus the group contributes to $4\cdot3\cdot3\beta[2\ (\Delta\vec{r}_2\cdot\vec{r}_3)^2]/8d^2 = 4\beta\gamma^2d^2$, and the combined contribution from these two groups is $6\beta\gamma^2d^2$.

Appendix C

Starting with Eq.(24) and using the SSHS model $\alpha = \alpha_I$, one finds the mixing enthalpy parameter Ω to be

$$\Omega = \frac{3}{2}\,\bar{\alpha}(d_{AC}-d_{BC})^2 N_o, \tag{C-1}$$

where N_o is Avagadro's number and $\bar{\alpha} = (1/2)\ (\alpha_{AC}+\alpha_{BC})$. Then relate the mean square bond length fluctuation $\langle\zeta^2\rangle$ at the melting temperature T_m to T_m for a compound by equating the average potential energy per unit cell to half of the thermal energy:

$$\langle P.E.\rangle \simeq 4\cdot\frac{3}{2}\alpha\langle\zeta^2\rangle = \frac{1}{2}(2\cdot 3kT_m) \tag{C-2}$$

where k is the Boltzmann constant. Defining a Liedermann ratio of melting χ_m by

$$(\langle\zeta^2\rangle)^{1/2} \approx \chi_m d \tag{C-3}$$

and choosing the mixing criterion to be $T_c/T_m < 1$, where T_m now is the smaller value of the two melting temperatures of the constituent compounds, we require that

$$\frac{T_c}{T_m} = \frac{\Omega k}{4R_g \alpha \chi_m^2 d^2} = \frac{3}{8} \frac{(d_{AC}-d_{BC})^2}{\chi_m^2 \, d^2} < 1, \qquad \text{C-4}$$

or

$$|\delta_o| / \delta_m < 1, \qquad \text{C-5}$$

where $\delta_m = 1.63\ \chi_m$ and $|\delta_o|$ is the percentage bond length difference.

REFERENCES

* Visiting professor at Stanford University.

1. J. C. Mikkelsen and J.B. Boyce, Phys. Rev. Lett., 49, 1412 (1982); Phys. Rev. B28, 7130 (1983).

2. A. Zunger and J. E. Jaffe, Phys. Rev. Lett. 51, 662 (1983).

3. J. L. Martins and A. Zunger, Phys. Rev. B30, 6217 (1984).

4. A. Balzarotti, A. Kisiel, N. Motta, M. Zimmal-Starnawska, M. T. Czyzyk, and M. Podgorny, Phys. Rev. B30, 2295 (1984).

5. A.-B. Chen and A. Sher, "Bonding Theory of Semiconductor Alloys," Microscience 3 (A multiclient publication of SRI International, Menlo Park, California, 1984).

6. A. Sher, A.-B. Chen, and W. E. Spicer, 13th International Conference on Defects in Semiconductors, August 12-17, 1984, Coronado, California, p.43.

7. W. A. Harrison, Electronic Structure and Properties of Solids (W. H. Freeman and Co., San Francisco, 1980); "The Bonding Properties of Semiconductors," Microscience, 3, p. 35 (A multiclient publication of

SRI International, Menlo Park, California, 1983); and Phys. Rev. B27, 3592 (1983).

8. K. Shih, W. E. Spicer, W. A. Harrison, and A. Sher, Phys. Rev. B (to appear in Jan 15, 1985).

9. A. Baldereschi and J. J. Hopfield, Phys. Rev. Lett. 28, 171 (1972).

10. R. M. Martin, Phys. Rev. B1, 4005 (1970).

11. The experimental elastic constants were taken from the tabulated values in Refs. 10 and 25 except AlP and AlAs, for which we used the estimated values by J. D. Wiley in Semiconductors and Semimetals, Vol. 10, edited by R. K. Willardson and A. C.Beer (Academic Press, New York, 1975), p. 134.

12. R. Kibuchi, Physica 103B, 41 (1981).

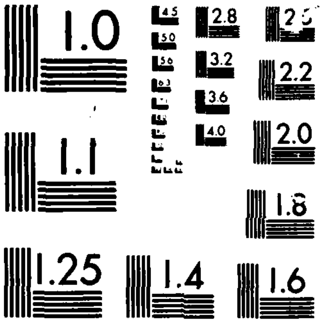
13. A.-B. Chen and A. Sher, unpublished.

14. J. B. Boyce and J. C. Mikkelsen, unpublished data for Zn(Se, Te). We thank them for their communication prior to publication.

15. Extrapolated results quoted in Ref. 3.

16. G. B. Stringfellow, J. Cryst. Growth 27, 21 (1974); 58, 194 (1982).

AD-A166 995   SEMICONDUCTOR ALLOY THEORY(U) AUBURN UNIV ALA DEPT OF   2/2
PHYSICS A CHEN 27 SEP 85 AFOSR-TR-86-0159
AFOSR-84-0282
UNCLASSIFIED   F/G 20/12   NL


1.0
4.5
5.0
5.6
6.3
2.8
3.2
3.6
4.0
2.5
2.2
2.0
1.1
1.8
1.25
1.4
1.6

MICROCOP CHART

17. J. A. Van Vechten, in Semiconductor Handbook, Vol. 3, edited by S. P. Keller (North-Holland, Amsterdam, 1980), p. 1

18. P. A. Fedders and M. W. Muller, J. Phys. Chem. Solids 45, 685 (1984).

19. M. B. Panish and Ilegems, Prog. Solid State Chem. 7, 39 (1972); A. Langier, Rev. Phys. Appl. 8, 2959 (1973); and Ref. 16.

20. T. P. Pearshall, R. Carles, and J. C. Portal, Appl. Phys. Lett. 42, 436 (1983).

21. K. Kakimoto and T. Katoda, Appl. Phys. Lett. 40, 826 (1982).

22. P. Parayanthal and F. H. Pollak, Phys. Rev. Lett. 52, 1922 (1984).

23. G. P. Strivastova, J. L. Martins, and A. Zunger, Phys. Rev. B (in press).

24. N. A. Groyunova, A. S. Barshchvevskii, and D. N. Fretiakov, in Semiconductors and Semimetals, Vol. 4, Edited by R. K. Willardson and A. C. Reer (Academic Press, New York, 1968), Ch. 1.

25. S. S. Mitra and N. E. Massa, in Handbook on Semiconductors, Vol. 1, edited by W. Paul (North-Holland, 1982), p. 29.

26. L. M. Foster, J. Electrochem. Soc. Solid State Sci. and Techn. 121, 1662 (1974).

27. L. M. Foster and J. F. Woods, J. Electrochem. Soc. Solid-State Sci. and Techn. 1919, 504 (1972).

28. R. Kubo, Statistical Mechanics (Univ. of Tokyo, 1971), p. 336.

29. The critical temperature will be reduced by a few percent if more elaborate models are used. See . W. Christian, Transformations in Metals and Alloys, 2nd. edition, edited by D. W. Hopkins (Pergamon Press, New York, 1981), p. 194.

30. American Institute of Physics Handbook, 3rd ed. (McGraw-Hill, New York, 1972), p. 9-16.

31. K. C. Hass, R. J. Lempert, and H. Ehrenreich, Phys. Rev. Lett. 52, 77 (1984).

END
FILMED

5-86

DTIC